INDEX TO VOLUME 63D

July to December 1959

A

Page

Agy, V., K. Davies, Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Airglow and aurora, origin of [OI] 5577 ... 15
Antenna gain, path, in an exponential atmosphere ... 273
Antennas, slotted-cylinder, pattern synthesis ... 303
Atmosphere, exponential reference, Central Radio Propagation Laboratory ... 287
Atmosphere, path antenna gain in an exponential ... 315
Atmosphere, upper, excitation mechanisms of the oxygen 5577 emission ... 273
Atmospheric horizontal inhomogeneity, effect of upon ray tracing ... 319
Aurora, a monochromatic low-latitude ... 297
Auroral and subauroral zones, [OI] 5577 in, comparison of absolute intensities ... 19

B

Bean, R. F., Climatology of ground-based radio ducts ... 29
—, *B. A. Cahoon*, Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
—, *J. D. Horn*, Radio-refractive-index climate near the ground ... 259
—, *L. P. Riggs, J. D. Horn*, Synoptic study of the vertical distribution of the radio refractive index ... 249
—, —, Synoptic variation of the radio refractive index ... 91
—, *G. D. Thayer*, Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Bremmer, H., Mode expansion in the low-frequency range for propagation through a curved stratified atmosphere ... 75
Broadcasting station, worldwide standard-frequency, power requirements and choice of an optimum frequency ... 35
Budden, K. G., Effect of small irregularities on the constitutive relations for the ionosphere ... 135

C

Cahoon, B. A., B. R. Bean, Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Climate, radio-refractive-index, near the ground ... 259
Climatology of ground-based radio ducts ... 29
Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
Conda, A. M., J. R. Wait, Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181
Correlation of solar noise fluctuations in harmonically related bands ... 293

D

Davies, K., V. Agy, Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Decker, M. T., Transmission and reflection by a parallel wire grid ... 87
Density, local electron, method for measuring from an artificial satellite ... 312
Devaney, T. E., A. N. Smith, Fields in electrically short ground systems ... 175
Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181

E

Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
Effect of small irregularities on the constitutive relations for the ionosphere ... 135
Electrically short ground systems, fields ... 175
Electromagnetic waves, diffraction of, by smooth obstacles for grazing angles ... 181
Electron density, local, method for measuring from an artificial satellite ... 325
Ellyett, C., J. M. Watts, Stratification in the lower ionosphere ... 117
Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319
Exponential atmosphere, path antenna gain ... 273
Exponential reference atmosphere, Central Radio Propagation Laboratory ... 315

F

Fields in electrically short ground systems ... 175
Fluctuations in harmonically related bands, correlation of solar noise ... 293
Frequency and power requirements for a worldwide standard-frequency broadcasting station ... 35

G

Gallet, R. M., R. A. Helliwell, Origin of "very-low-frequency emissions" ... 21
Gates, D. M., Preliminary results of the National Bureau of Standards radio and ionospheric observations during the International Geophysical Year ... 1
Grazing angles, diffraction of electromagnetic waves by smooth obstacles ... 181
Grazing angles, study at 1,046 megacycles per second of the reflection coefficient of irregular terrain ... 235
Grid, parallel wire, transmission and reflection ... 87

Page

Ground-based radio ducts, climatology ... 29
Ground systems, electrically short, fields ... 175

H

Hartman, W. J., R. E. Wilkerson, Path antenna gain in an exponential atmosphere ... 273
Helliwell, R. A., R. M. Gallet, Origin of "very-low-frequency emissions" ... 21
Horn, J. T., B. R. Bean, Radio refractive-index climate near the ground ... 259
—, —, *L. P. Riggs*, Synoptic study of the vertical distribution of the radio refractive index ... 349
Householder, J., J. R. Wait, Pattern synthesis for slotted-cylinder antennas ... 303

I

Index, radio refractive, synoptic study of the vertical distribution ... 249
Intensities, absolute, of [OI] 5577 in the auroral and subauroral zones, comparison ... 19
International Geophysical Year, radio and ionospheric observations ... 1
Ionosphere, effect of small irregularities on the constitutive relations ... 135
Ionosphere, lower, stratification in ... 117
Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Irregularities, small, effect of on the constitutive relations for the ionosphere ... 135
Irregularities, tropospheric, radio-wave scattering ... 205

J

Janes, H. B., M. C. Thompson, Jr., Measurements of phase stability over a low-level tropospheric path ... 45
Jean, A. G., W. L. Taylor, Very-low-frequency radiation spectra of lightning discharges ... 199

L

Lightning discharges, vlf radiation spectra of ... 199
Low-frequency propagation paths in arctic areas ... 99
Low-latitude aurora, a monochromatic ... 297

M

Maloney, L. J., R. E. McGavin, Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Marovich, E., F. E. Roach, A monochromatic low-latitude aurora ... 297
—, —, *J. W. McCaulley*, Origin of [OI] 5577 in the airglow and the aurora ... 15
Maxwell, E. L., A. D. Watt, Low-frequency propagation paths in arctic areas ... 99
McCaulley, J. W., F. E. Roach, E. Marovich, Origin of [OI] 5577 in the airglow and the aurora ... 15
—, —, *C. M. Purdy*, Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
McGavin, R. E., L. J. Maloney, Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Measurements of phase stability over a low-level tropospheric path ... 45
Method for measuring local electron density from an artificial satellite ... 325
Mode expansion in the low-frequency range for propagation through a curved stratified atmosphere ... 75
Monochromatic low-latitude aurora ... 297

N

Norton, K. A., System loss in radio wave propagation ... 53

O

Origin of [OI] 5577 in the airglow and the aurora ... 15
Origin of "very-low-frequency emissions" ... 21
Oxygen 5577 emission in the upper atmosphere, excitation mechanisms ... 319

P

Path antenna gain in an exponential atmosphere ... 273
Pattern synthesis for slotted-cylinder antennas ... 303
Phase stability, measurements of, over a low-level tropospheric path ... 45
Plush, R. W., A. D. Watt, Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 35
Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 35
Preliminary results of the National Bureau of Standards radio and ionospheric observations during the International Geophysical Year ... 1
Propagation, low-frequency, in arctic areas ... 99
Propagation through a curved stratified atmosphere, mode expansion in the low-frequency range ... 75
Purdy, C. M., F. E. Roach, J. W. McCaulley, Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19

R

Page

Radiation spectra, vlf, of lightning discharges ... 199
Radio and ionospheric observations during the International Geophysical Year ... 1
Radio ducts, ground-based, climatology ... 29
Radio-refractive-index climate near the ground ... 259
Radio refractive index, synoptic study of the vertical distribution ... 249
Radio refractive index, synoptic variation ... 91
Radio-wave propagation, system loss ... 53
Radio-wave scattering by tropospheric irregularities ... 205
Ray tracing, effect of atmospheric horizontal inhomogeneity ... 287
Reflection and transmission by a parallel wire grid ... 87
Reflection coefficient of irregular terrain at grazing angles, study at 1,046 megacycles per second ... 235
Refractive index, radio, synoptic variation ... 91
Riggs, L. P., B. R. Bean, Synoptic variation of the radio refractive index ... 91
—, —, *J. D. Horn,* Synoptic study of the vertical distribution of the radio refractive index ... 249
Roach, F. E., E. Marovich, A monochromatic low-latitude aurora ... 297
—, *J. W. McCaulley, E. Marovich,* Origin of [OI] 5577 in the airglow and the aurora ... 15
—, —, *C. M. Purdy,* Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
—, *E. Tandberg-Hanssen,* Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319

S

Satellite, artificial, method for measuring local electron density ... 325
Slotted-cylinder antennas, pattern synthesis ... 303
Smith, A. N., T. E. Devaney, Fields in electrically short ground systems ... 175
Smooth-obstacle diffraction of electromagnetic waves for grazing angles ... 181
Spectra, vlf radiation, of lightning discharges ... 199
Stability, phase, measurements over a low-level tropospheric path ... 45
Standard-frequency broadcasting station, worldwide, the power requirements and choice of an optimum frequency ... 35
Storey, L. R. O., Correlation of solar noise fluctuations in harmonically related bands ... 293
—, Method for measuring local electron density from an artificial satellite ... 325
Stratification in the lower ionosphere ... 117
Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Sweep-frequency pulse technique at oblique incidence, ionospheric investigations using ... 151
Synoptic study of the vertical distribution of the radio refractive index ... 249
Synoptic variation of the radio refractive index ... 91
Synthesis for slotted-cylinder antennas ... 303
System loss in radio wave propagation ... 53

T

Tandberg-Hanssen, E., F. E. Roach, Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319
Taylor, W. L., A. G. Jean, Very-low-frequency radiation spectra of lightning discharges ... 199
Terrain, irregular, study at 1,046 megacycles per second of the reflection coefficient at grazing angles ... 235
Thayer, G. D., B. R. Bean, Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Thompson, M. C., Jr., H. B. Janes, Measurements of phase stability over a low-level tropospheric path ... 45
Transmission and reflection by a parallel wire grid ... 87
Tropospheric irregularities, radio-wave scattering ... 205
Tropospheric path, low-level, measurements of phase stability ... 45

V

Variation, synoptic, of the radio refractive index ... 91
Very-low-frequency emissions, origin of ... 21
Very-low-frequency radiation spectra of lightning discharges ... 199

W

Wait, J. R., A. M. Conda, Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181
—, *J. Householder,* Pattern synthesis for slotted-cylinder antennas ... 303
Watt, A. D., E. L. Maxwell, Low-frequency propagation paths in arctic areas ... 99
—, *R. W. Plush,* Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 3
Watts, J. M., C. Ellyett, Stratification in the lower ionosphere ... 117
Wheelon, A. D., Radio-wave scattering by tropospheric irregularities ... 205
Wilkerson, R. E., W. J. Hartman, Path antenna gain in an exponential atmosphere ... 273
Worldwide standard-frequency broadcasting station, the power requirements and choice of an optimum frequency ... 35

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Preliminary Results of the National Bureau of Standards Radio and Ionospheric Observations During the International Geophysical Year

David M. Gates

(March 2, 1959)

A review is given of the activities of the National Bureau of Standards during the International Geophysical Year. The equipment used on each project is described and preliminary results of the observations are given. The following areas of research are discussed: (1) World Warning Agency, (2) Ionospheric Vertical Sounding Stations, (3) VHF Propagation, (4) VHF Equatorial Forward Scatter, (5) Radio Noise Network, (6) Radio Satellite Observations, (7) Airglow Observations, and (8) World Data Center for Airglow and Ionosphere.

1. Introduction

The International Geophysical Year program was the greatest single international collaborative research effort ever devised by man to probe the properties of the planet earth from the depths of the oceans to the domain of interplanetary space. As a part of this program, the scientists of the Central Radio Propagation Laboratory, of the National Bureau of Standards in Boulder, Colorado, undertook a vast expansion of their radio research efforts in order to contribute to this international program and to take advantage of the many different simultaneous measurements being made throughout the world. Although most of the efforts were of a worldwide synoptic character, it was important that certain selected specialized experiments be carried out to coincide with the overall effort. The NBS experimental program, during the IGY, ranged from the large synoptic network of ionospheric vertical soundings stations to selected experiments concerned with equatorial forward scatter, sporadic-*E*, radio noise, night airglow and observations on satellite signals. This article is a brief summary of some of these projects at NBS and a preliminary report concerning some of the results available at this time.

Since the IGY was selected to coincide with a period of maximum activity in the solar atmosphere, it was essential that a close watch on the sun be maintained throughout this period. The solar patrols consisted of visual, optical, photographic, photometric, and radio observations of the sun throughout every minute of the IGY period. The IGY World Warning Agency headed by R. C. Moore at the NBS station at Fort Belvoir, Virginia, was designated to process and filter these data, to maintain a running account of the "state of the sun", and to alert the scientists of the world to periods of unusual solar activity. At times, during those Alerts, Special World Intervals were declared thereby accelerating scientific observations throughout the world. During the entire IGY period 21 Special World Intervals (SWI's) totaling 45 days were declared. During 16 of these SWI's 18 geomagnetic disturbances occurred; two of the disturbances being followed by second storms. Only four SWI's totaling 7 days were not followed by disturbances and one interval of 1-day duration was initiated after the storm had started. Approximately two-thirds of the severe geomagnetic disturbances were predicted and one-third were missed. However, previous to the IGY it was thought that if two major storms were forecast during the entire IGY period this would represent adequate success.

2. IGY Ionospheric Vertical Sounding Observations by NBS

During the International Geophysical Year a large network of vertical incidence ionospheric sounding stations (about 150) was placed in operation throughout the world. Since the vertical incidence pulse technique is one of the most reliable methods for obtaining a measure of the electron density and the electron density profile of the ionosphere, it was important to establish stations of this nature at as many strategic locations as possible. Prior to the IGY the National Bureau of Standards operated 20 stations throughout the world, while during the IGY they operated directly or cooperated with other agencies in the operation of 34 stations. Table 1 lists these stations, gives the geographic and geomagnetic coordinates, the angle of magnetic dip, and lists the equipment used at each station. All stations were on a 15-min basic-sounding schedule. Much of the reduction of the data was done at the stations, but all data sent to Boulder on a monthly or bimonthly basis was reviewed in order to assure smooth and consistent operation of the program. Engineering teams were periodically sent to the stations in order to install any modification of the equipment and to assist where needed. The vertical incidence program at the National Bureau of Standards has been under the direction of A. H. Shapley, R. W. Knecht, and H. G. Sellery.

TABLE 1. *Vertical sounding stations operated as part of the NBS observational program during the IGY*

Station	Equipment	Geographic coordinates		Geomagnetic coordinates		Dip angle	Time used	Operated by
		lat.	long.	lat.	long.			
Adak, Alaska	C-3	N51 54	W176 39	N47.3	E240.0	64	180W	Signal Corps.
Anchorage, Alaska	C-3	N61 14	W149 53	N60.9	E258.1	74	150W	NBS.
Baguio, Phil. Is.	C-2	N16 25	E120 36	N51.1	E189.3	18	120E	Manila Observatory
Barrow, Alaska	C-2	N71 08	W156 48	N68.4	E241.2	81	150W	NBS.
Bogota, Colombia	C-4	N4 32	W74 15	N15.9	E354.5	32	75W	Geophys. Inst. Colombian Andes.
Byrd, Antarctica	C-3	S80	W120	S70 36	E336 01	−75	120W	NBS.
Cape Hallett, Antarctica	C-4	S72 25	E170 55	S74.6	E279.0	−84	165E	New Zealand.
Chiclayo, Peru	C-3	S6 48	W79 49	N4.4	E349.2	10	75W	Geophys. Inst. Huancayo.
Chimbote, Peru	C-4	S9 04	W78 35	N2.2	E350.4	6	75W	Geophys. Inst. Huancayo.
Concepcion, Chile	C-4	S36 35	W72 59	S25.3	E356.5	−35	75W	Univ. of Concepcion.
Ellsworth, Antarctica	C-4	S77 43	W41 08	S66 56	E14 42	−68	45W	NBS.
Ft. Belvoir, Va.	C-3	N38 44	W77 08	N50.1	E350.1	71	75W	NBS.
Ft. Monmouth, N.J.	C-4	N40 23	W74 08	N51.7	E353.9	71.5	75W	Signal Corps.
Fairbanks, Alaska	C-3	N64 51	W147 50	N64.7	E256.5	77	150W	Univ. of Alaska.
Godhavn, Greenland	C-3	N69 16	W53 30	N79.9	E32.5	81	45W	Danish URSI Comm.
Grand Bahama Is., B.W.I.	C-4	N26 40	W78 22	N37.9	E349.3	59	75W	Signal Corps.
Huancayo, Peru	C-4	S12 03	W75 20	S0.6	E353.8	2	75W	Geophys. Inst. Huancayo.
Fletcher's Ice Island	C-2	N85	W100	(Drifting)		88	75W	Signal Corps.
La Paz, Bolivia	C-4	S16 29	W68 03	S5.0	E0.9	−5	60W	University of San Andres.
Little America, Ant.	C-4	S78 16	W162 28	S74 08	E312 04	−80.5	165W	NBS.
Maui, T.H.	C-4	N20 50	W156 30	N20.8	E268.1	38	150W	NBS.
Natal, Brazil	C-4	S5 20	W35 07	N4.2	E33.9	−2	30W	Brazilian Navy.
Narsarssuak, Greenland	C-3	N61 11	W45 25	N71.2	E37.6	77	45W	Danish URSI Comm.
Okinawa, Ryu Kyu I.	C-3	N26 19	E127 47	N15.5	E195.7	37	135E	Signal Corps.
Panama Canal Zone	C-3	N9 23	W79 53	N20.6	E348.5	39	75W	NBS.
Puerto Rico	C-3	N18 30	W67 12	N30.0	E2.0	51	60W	NBS.
Reykjavik, Iceland	C-2	N64 08	W21 47	N70.2	E69.8	74	15W	Iceland Post & Telegraph.
San Francisco, Calif.	C-3	N37 26	W122 10	N43.7	E298.4	62	120W	Stanford University.
South Pole Station	C-3	S90		S78 30	0	−75	0	NBS.
St. Johns, Newfoundland	C-3	N47 37	W52 42	N58.5	E21.4	72.5	60W	Signal Corps.
Talara, Peru	C-3	S4 38	W81 18	N6.6	E347.7	13	75W	Geophys. Inst. Huancayo.
Thule, Greenland	C-4	N76 33	W68 50	N88 03	E1 08	86	75W	Signal Corps.
White Sands, N. Mex.	C-3	N32 24	W106 52	N41.2	E316.4	60	105W	Signal Corps.
Wilkes, Antarctica	C-4	S66 25	E110 27	S77.9	E178.8	−81	105E	NBS.

2.1. Equatorial Close-Spaced Chain

Interesting results have come from the ionospheric soundings from the equatorial close-spaced chain of stations in South America [1].[1] These stations were located at Talara, Chiclayo, Chimbote, Huancayo, and La Paz. Their geographic and geomagnetic coordinates are given in table 1. The hourly ionograms for 15 October 1957 were reduced to true height profiles of electron density using the Budden matrix method [2]. In figure 1 the variation of electron density at fixed heights is shown for the 24-hr period at Huancayo. These curves are typical of the other four stations as well. The sharp increase at sunrise in the electron density should be noted as well as the fact that the intermediate levels attain a maximum electron density earlier than the higher levels. In addition, the lower levels exhibit a Chapman-like behavior, i.e., electron density proportional to the square root of the cosine of the sun's zenith angle. Particularly notable is the abrupt rise in height of the electron density at sunset, an effect considerably larger than expected from straight recombination alone.

Equatorial sporadic *E*, which occurs during most daylight hours throughout the year at stations near the equator, appears on the ionogram as a diffuse trace often extending 20 km or more in height and seldom blanketing the *F* region echoes. The extent of the equatorial sporadic *E* can readily be ascertained from figure 2 which shows the percentage occurrence as a function of the magnetic dip of the station. Eleven months of data (July 1957 through May 1958) were averaged to obtain the percentages for Talara, Chiclayo, Chimbote, and Huancayo, however, only one month of data was available for La Paz. From the curve it is clear that this form of sporadic *E* is a very distinctive equatorial phenomena, and that it occurs with a frequency greater than 35 percent only within a belt spread about 5.5° to either side of the geomagnetic equator and has dropped to a frequency of only 5 percent by 8° from

[1] Figures in brackets indicate the literature references at the end of this paper.

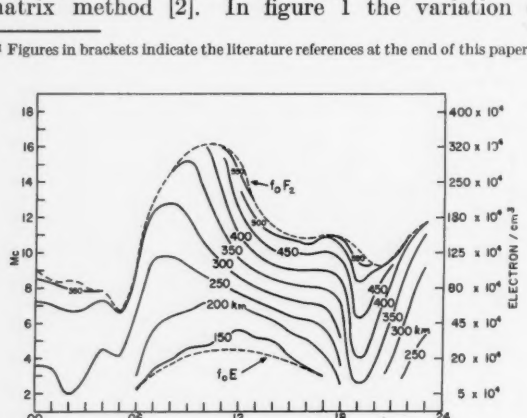


FIGURE 1. *Electron density at fixed heights for huancayo.* October 15, 1957.

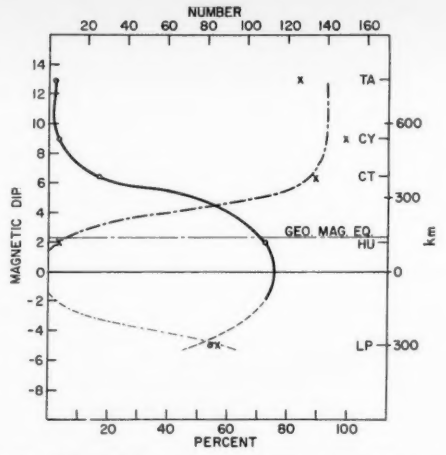


FIGURE 2. *Percentage occurrence of* Es-q *along the equatorial close-spaced chain, 06–(———), and total number of hourly soundings showing prominent blanketing by* Es, *i.e.* fbEs$>$ *5 Mc (——— — ———), July 1957 through May 1958. TA-Talara, CY-Chiclayo, CT-Chimbote Hu-Huancayo, LP-DaPaz.*

July 1957 through May 1958. TA = Talara, CY = Chiclayo, CT = Chimbote' HU = Huancayo, and LP = LaPaz.

the geomagnetic equator. Although earlier workers [3, 4] had suggested a somewhat wider belt for equatorial sporadic E, recent IGY experiments on the earth's magnetic field in the vicinity of the geomagnetic equator by S. E. Forbush, as recently reported by H. Odishaw [5], give a similar width to the equatorial electrojet. Note that the scale along the right-hand vertical axis represents distance perpendicular to the contours of equal dip and not distance between stations.

Although the equatorial sporadic E persisted throughout most days of the year, occasionally it would disappear suddenly for periods of minutes to hours in the early afternoon. Matsushita [6] has established a definite lunar dependence for this effect, which suggested that a similar relationship might hold true for the time of first appearance for equatorial sporadic E, since these times varied from as early as 0600 and as late as 1000. In contrast, the late afternoon disappearance occurred regularly between 1730 and 1830. In order to test this hypothesis the data for Huancayo from May 1957 to April 1958 were divided into the three seasons (equinox and June and December solstice). Three day running means of the times of first appearance were computed and plotted against the lunar phase as shown in figure 3. A lunar correlation is clearly evident in the equinoctial and December-solstice curves with equatorial sporadic E appearing earlier in the morning on days of new and full moon. The data for the June solstice do not suggest quite as strong a lunar control, however, as more data become available the scatter of the points may be reduced. Matsushita had explained the lunar effect for the midday disappearance in terms of ionospheric electric currents. He assumed that the equatorial sporadic E was caused by the vertical drifting of

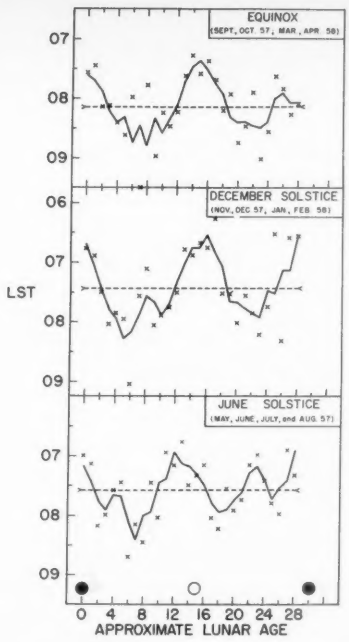


FIGURE 3. *Lunar influence on time of first appearance of* Es-q *at Huancayo.*

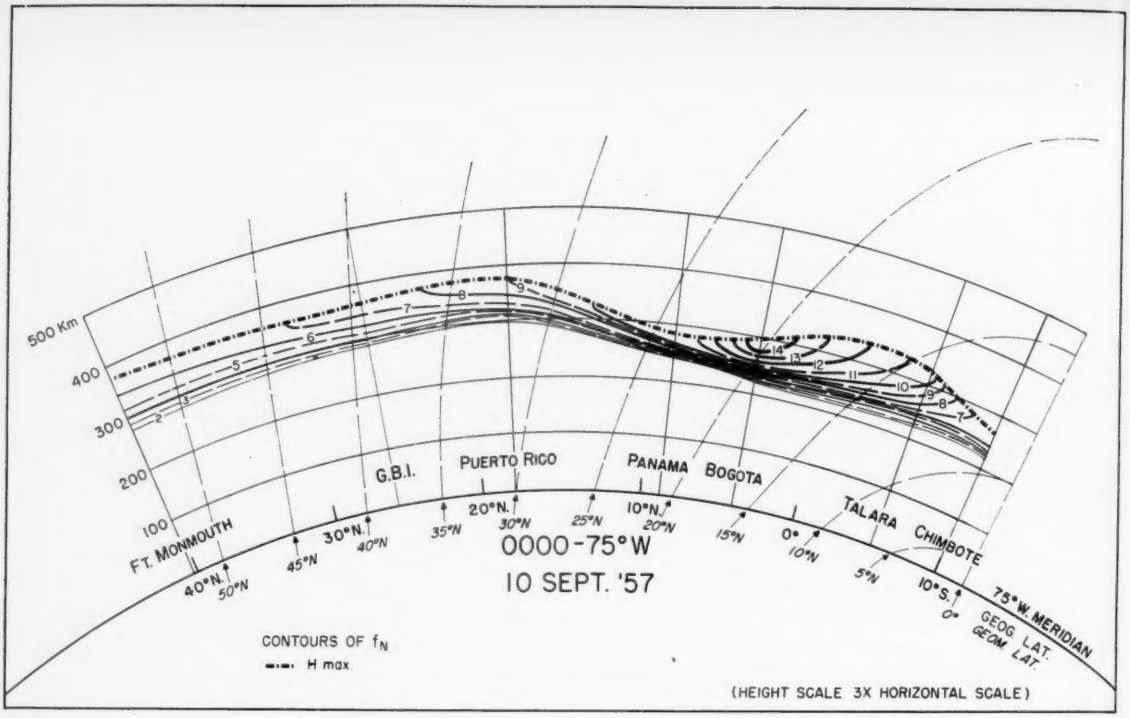
Three-day running mean plotted.

electrons in the eastward-flowing electric current and that the lunar tides induce a westward-flowing counter-current which reduces the strength of the sporadic E. Pursuing these arguments further, Knecht [7] of the Central Radio Propagation Laboratory points out that during the new and full moon the lunar-induced current adds an easterly flow to the already established solar induced easterly during the early morning hours, whereas during the first and third quarter the lunar tide counteracts the easterly solar current thus retarding the appearance of the sporadic E in the morning. It would appear from these observations that Matsushita's hypothesis concerning the electron-current systems has been further verified.

2.2. Electron Density Profiles along the 75° W Meridian

A true height electron density analysis has been made for the following stations located near the 75° W meridian: Chimbote and Talara, Bogota, Panama, Puerto Rico, Grand Bahama Island, and Fort Monmouth, New Jersey. The Budden matrix method was used with the assistance of the IBM 650 computer to reduce the ionogram data. Another computer program was used to interpolate the values of electron density at fixed heights, compute the height of the maximum and integrate the electron density up to h_{max}. This research was carried out at NBS by J. W. Wright and G. H. Stonehocker.

The hourly ionograms from the seven stations were reduced to give true heights of reflection at frequency intervals of 0.2 Mc from 1.8 Mc to f_0F_2.



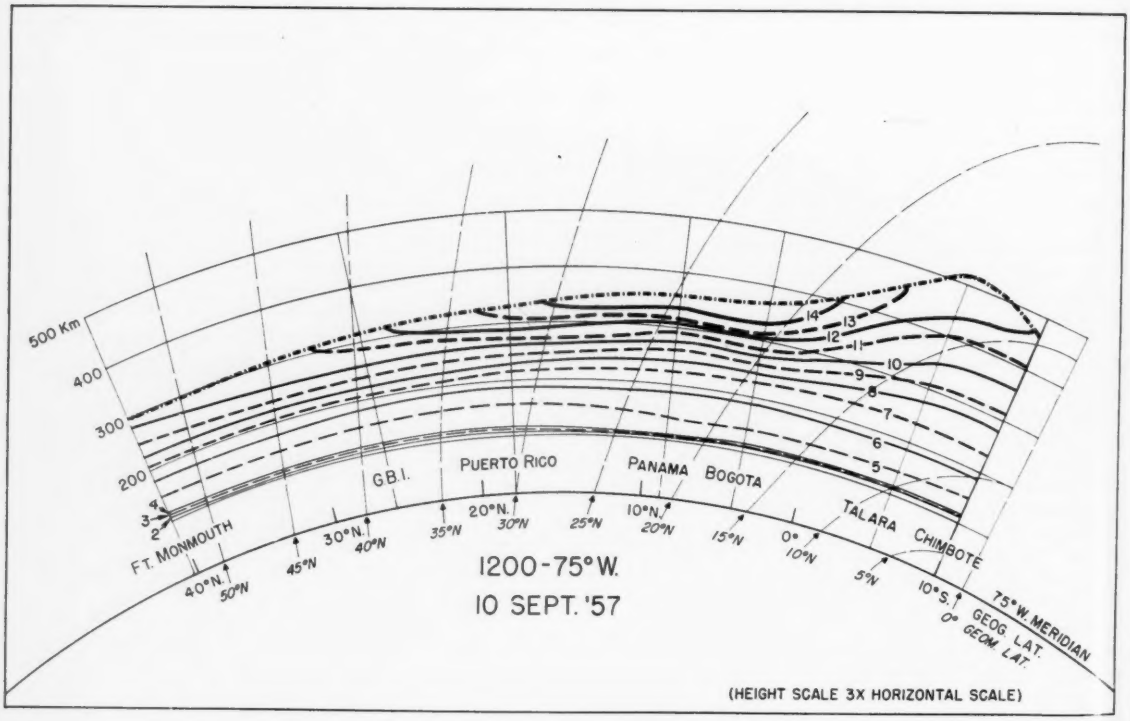
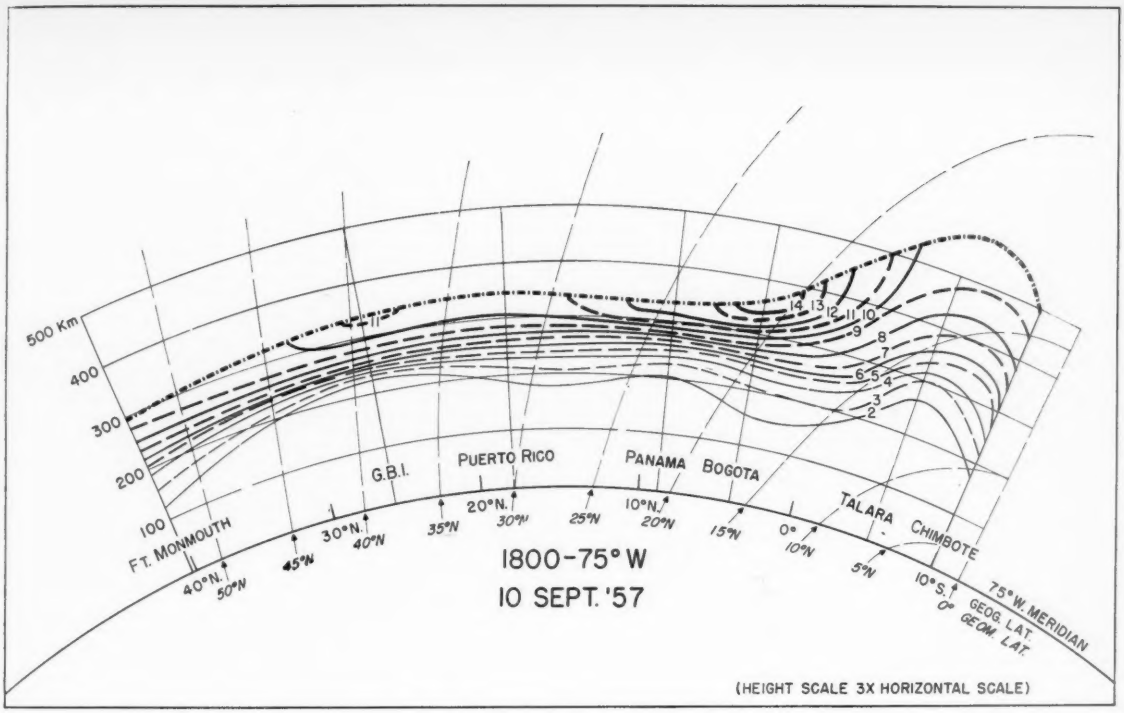
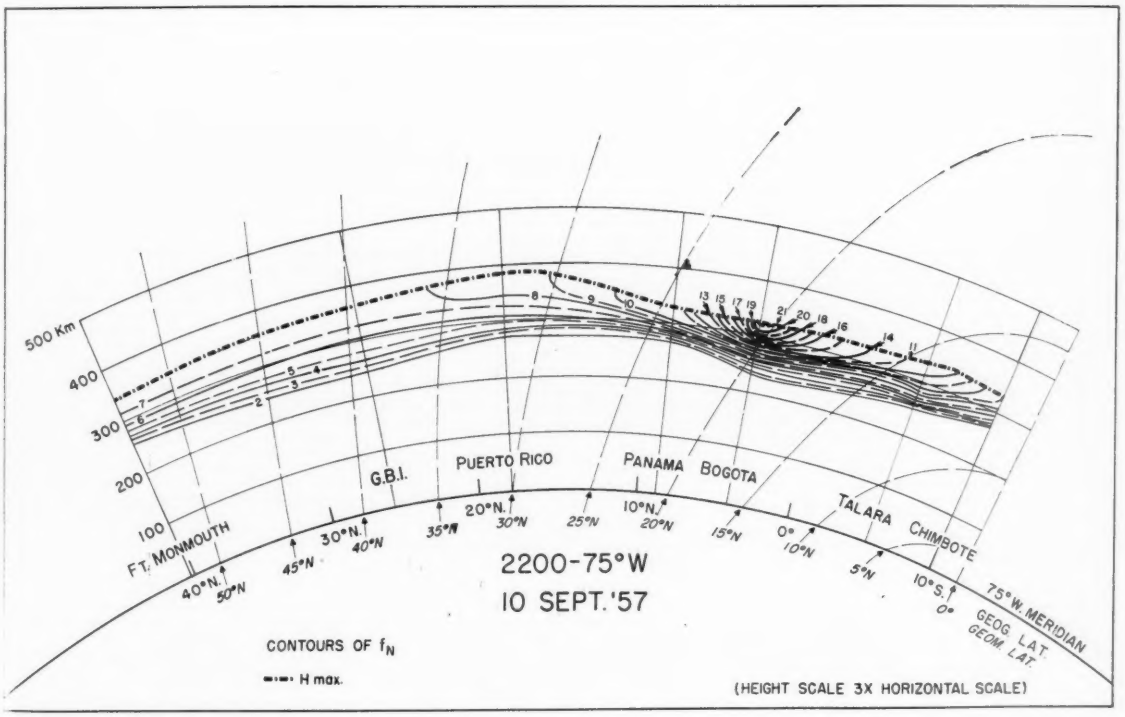


FIGURE 4. *True height electron density profiles along the 75°*





the 75° W meridian, between 40° N and 10° S.

A sample of the results, after interpolation, is shown in figure 4 for 0000, 1200, 1800, and 2200 universal time (u.t.). Here contours of f_N, the plasma frequency, are shown on a height versus latitude plot. Similar plots have been made for every 2 hr throughout the 24-hr period, but these suffice to illustrate the phenomena. Several of the usual ionospheric features show up in these diagrams, but the new and interesting feature concerns the increase in electron density over Bogota during the late evening hours and the bulge in the ionosphere which developes over the equator. It is possible that this bulge may be the result of some form of tidal action but at the moment it is not fully understood.

2.3. Vertical Incidence Soundings from Antarctica

For the first time during the IGY there was established on the Antarctic continent a sizeable network of ionospheric vertical incidence stations [8]. Of a total of approximately a dozen, the following five were operated by the U.S.: Pole, Byrd, Little America, Ellsworth, and Wilkes. The data reported here represent the first from high-latitude southern stations and the first ever obtained at either the North or South Pole. Prior to making these observations there was considerable question concerning the state of the ionosphere above the pole, whether the ionosphere could exist throughout the 6-month period of darkness, whether diurnal variations existed in the polar winter ionosphere, and what the nature of the f_0F_2 variations over the Antarctic continent were like. The present data, representing the period June 1957 to May 1958, definitely supply the answers to some of these questions. Further data through January 1959 are now available but only the analysis of the earlier data will be presented here.

In figure 5 are shown, for five of the antarctic stations, the monthly median values of f_0F_2 plotted against time of day for each hour. For the Wilkes and Ellsworth stations, the normal temperate latitude behavior is seen with the diurnal variation reaching a maximum around midday. For the Little America and Byrd bases the maximum is reached during the evening hours and for the Pole there appears to be a weak diurnal variation with the maximum occurring in the early morning hours. It is curious indeed that the ionosphere over the pole in July has minimum f_0F_2 values which are greater than those over any other station. This is particularly surprising since throughout July the atmosphere above the pole is not illuminated below an altitude of about 450 km. The summertime (December) f_0F_2 values exhibit relatively little diurnal variation as expected.

Typical winter and summer ionograms taken at the Pole station are shown in figure 6. The winter-time-spread echo shown here is typical of magnet-

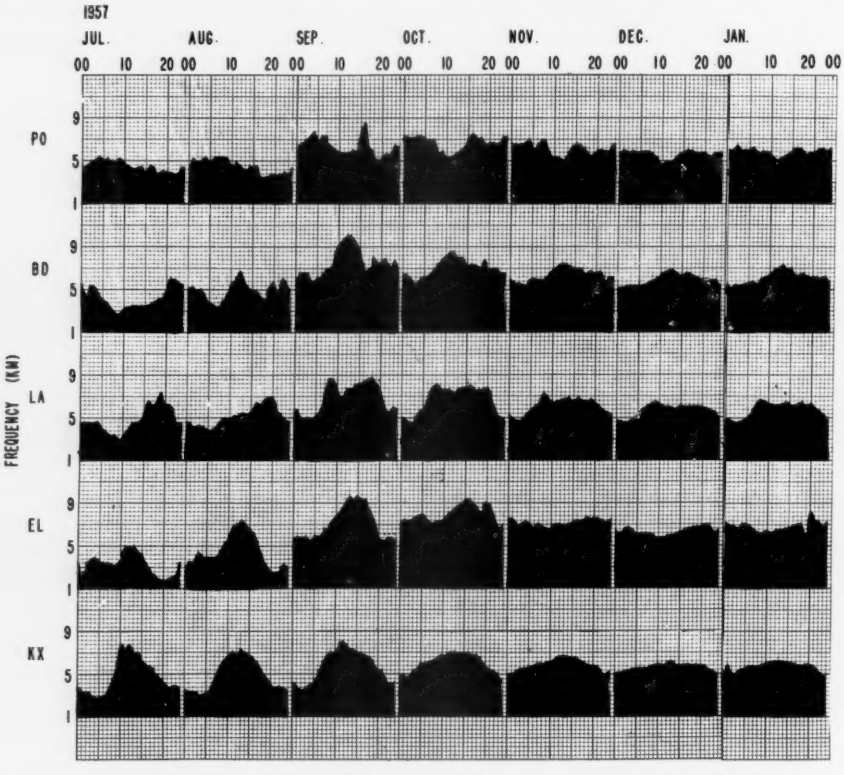


FIGURE 5. *The monthly median f_0F_2 for five antarctic stations.*
PO=Pole, BD=Bird, LA=Little America, EL=Ellsworth, KX=Wilkes.

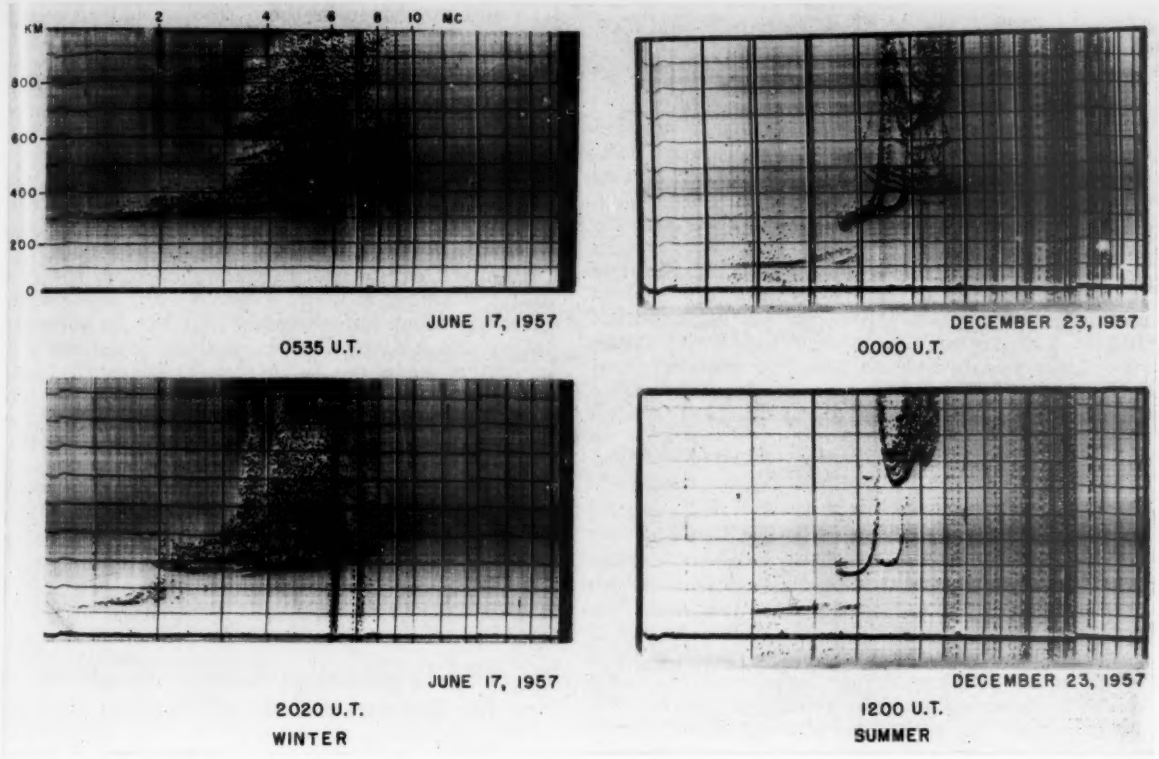


FIGURE 6. *Typical winter and summer ionograms taken at the South Geographic Pole.*

ically-undisturbed winter days. These soundings show that there is considerable ionization present as low as 250 km while the sun's rays do not strike below 450 km for July. Note also the appearance of the "night E" on the 2020 ionogram. The summertime ionograms shown on the right are typical of this season at the Pole. Particularly noteworthy is the constancy throughout the day of the critical frequency for E, F_1, and F_2.

A more detailed picture of the monthly median values of f_0F_2 for each hour of the day for the period June 1957 through May 1958 is shown in figure 7. The winter months (June, July, August) have the lowest f_0F_2 values and appear to have a 24-hr wave, whereas the equinox and summer values are higher and exhibit a 12-hr variation. A harmonic analysis of the monthly median values of f_0F_2 for these data is shown in figure 8 for the 24-, 12-, 8- and 6-hr components. The 24-hr component shows the winter values occurring at 0600 u.t. to be 90° out of phase with the summer values at approximately 0000 u.t. Magnetic noon (the time when the sun is in the meridian plane containing the geomagnetic pole and the geographic pole) occurs at 0500 u.t. thus suggesting that the F-layer density in the darkened part tends to reach a maximum during those hours when the auroral zone has the maximum amount of sunlight for that time of the year. The 12-hr wave appears at 0400 and 1600 u.t. and the 8-hr wave at 0730, 1530, and 2330 u.t. The 6-hr wave appears indefinite.

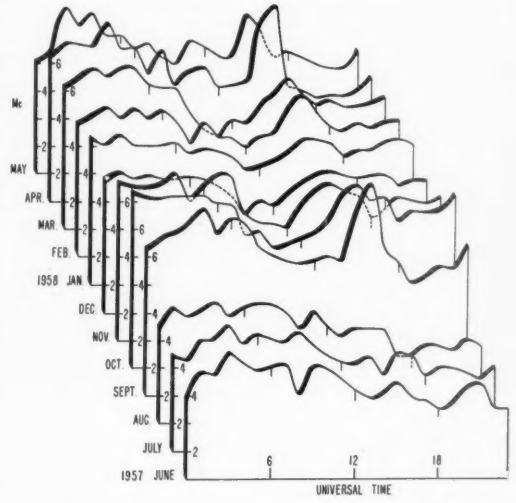


FIGURE 7. *Monthly median* f_0F_2 *for the South Pole Station.*

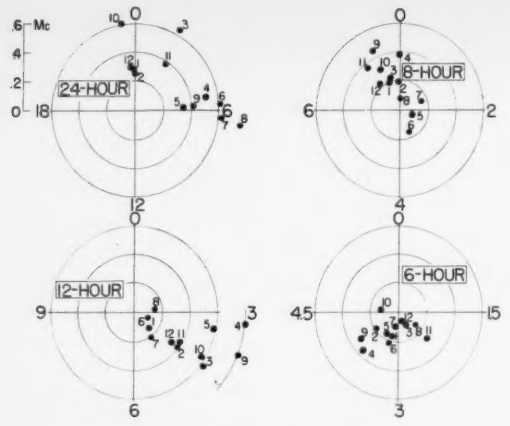


FIGURE 8. *Harmonic analysis of f_0F_2 for the South Pole.* June 1957 through May 1958.

Although the analysis is still in progress it would appear that a significant diurnal variation occurs in the f_0F_1 and f_0E during the polar summer. The principal wave seems to have a 24-hr period showing a maximum at 2200 for f_0F_1 and at 1300 for f_0E. Even more striking though is the summertime, daylight behavior of f_0F_2, f_0F_1, and f_0E monthly mean values for the Pole station when plotted against Zürich sunspot number as shown in figure 9.

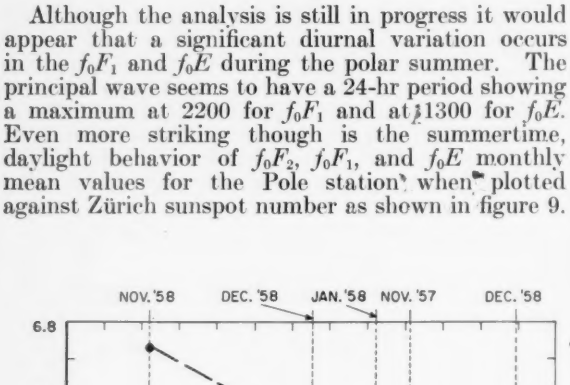


FIGURE 9. *Means of the monthly medians for summer months at the Pole Station as a function of the monthly Zürich sunspot numbers.*

It would appear from these results that over the Pole the F_2 region of the ionosphere becomes weaker, the E region strengthens, and the F_1 region is relatively insensitive to increased solar activity. This occurs at the time of the year when the polar ionosphere is completely and continuously illuminated. By comparison, it is well known that at stations throughout the world, the critical frequency for each region of the ionosphere increases with increasing sunspot number, since the increased solar activity produces an increase in ionizing radiation. Certainly it can be said that the data from the Pole station is unique in many respects. It raises several questions which unfortunately may not be answered until a much more detailed analysis is carried out. In certain instances further data will need to be obtained from the Pole station and careful comparison made with data from other Antarctic stations. Fortunately, the scientific observational program is continuing, although at a reduced level, in the Antarctica during 1959. Ionospheric soundings are being made at the Pole Station and at Byrd Base and it would be hoped that these measurements could be continued through the next minimum of solar activity, if not indefinitely.

3. VHF Propagation Comparison Between Far Eastern and American Circuits

From VHF transmissions over circuits in the Far East there had been evidence of more enhancements due to sporadic-E propagation than is observed in comparable latitudes in the rest of the world [3]. At times the signals were some 20 db or more higher than over circuits in the United States. Usually propagation at VHF frequencies takes place by scatter from the lower ionosphere and occasionally by sporadic-E enhancement. Because of the importance of sporadic E to communications and because of the great difficulty scientists have had in understanding the mechanisms creating sporadic E in the ionosphere it was particularly important to obtain detailed observations concerning this phenomenon in the Far East during the IGY. For this reason Dr. E. K. Smith, the project director, located one circuit in the Far East and a comparison circuit in the western hemisphere, each being a 2-kw narrow-band CW system operating over approximately an 800-mile path at 50 Mc.

The Far Eastern circuit extended between the Voice of America relay stations at San Fernando (Poro Point), La Union, Luzon, Philippine Islands and Onna, Okinawa, on 49.84 Mc. An additional transmitter located at Okuma, Okinawa, produced 49.68-Mc signals which were received by the Japanese Radio Research Laboratories at Hiraiso and Yonago, Japan. In the western hemisphere a similar circuit was set up on 49.76 Mc between the CRPL field station at Margarita, Panama Canal Zone, and the U.S. Naval Base at Guantanamo Bay, Cuba. Late

in the IGY an additional circuit was installed between the Canal Zone and Guayaquil, Ecuador, in cooperation with the Equatorial Scatter program. There was also in existence the circuit between Havana, Illinois and Boulder, Colo., operating at frequencies of 30.010, 49.605, and 73.88 Mc. This latter circuit has been a good comparison circuit for the others. All stations had 5-element Yagi antennas.

The first test measurements were obtained during September 1957 on both the Panama to Guantanamo circuit and the Poro Point to Onna circuit and both circuits became officially operative on 1 October 1957. The Havana, Illinois to Boulder, Colorado circuit became operative by November 1957. Essentially two experiments were performed during the IGY period; the regular sporadic-E type of observations and a special pulsed F-layer scatter experiment.

The expected fact that sporadic E is more intense in the Far East than in the Americas has been borne out. From the statistics gathered during the year sporadic E of moderate intensity would appear to be approximately 2 to 3 times as frequent between the Philippines and Okinawa as between Panama and Cuba or Panama and Ecuador.

As the observations progressed it was realized that the considerable enhancements taking place during evening hours might be due to scatter from the F region. At 50 Mc the signal often rose 30 to 40 db above the normal scatter signal level. During the IGY period the signal enhancement at 49.84 Mc showed an autumnal high with intermittent enhancement during the remainder of the year and a minimum during the vernal equinox. During the year, May 1956 to May 1957, over a circuit between the Philippines and Okinawa at 36.4 Mc, the evening anomaly showed a maximum during each equinoctial period. In order to verify the F-region origin for these anomalous signals a series of pulse measurements, giving range-time records, were carried out over the Philippines to Okinawa path during September 1958. Careful independent time standards were used to control the pulse repetition rate at the transmitting terminal and the oscilloscope sweep rate at the receiving terminal. The pulses were of 50-μsec duration, at a repetition rate 100/sec, with a peak power of about 2 kw. When received, the pulses were spread over nearly a millisecond. The F-region scattered signals were readily distinguished from the normal E-region echo and showed a delay time relative to the E-region signal between 40 and 1200 μsec. Seldom were sporadic-E signals between Panama and Cuba as strong as those in the Far East, despite the fact that in the Caribbean area the circuits spanned equivalent geomagnetic and geographic latitudes to those in the Far East.

A more detailed report of this Far Eastern VHF anomaly has been published [9]. These authors conclude that these anomalous signals are associated with low-latitude spread F as often recorded by ionosphere sounders and may indicate the presence of field-alined blobs of ionization in the F region.

4. VHF Equatorial Forward Scatter Experiments

Heretofore, experiments in VHF ionospheric forward scatter have been primarily conducted in north temperate latitudes and in the auroral zone, and not in the vicinity of the magnetic equator, where the ionosphere behaves differently and in a manner little understood. In this region are exhibited such phenomena as equatorial sporadic E, equatorial spread F, and a strong daytime current-system girdling the earth at the 100-km level, known as the "equatorial electrojet". Amateur radio communications spanning the magnetic equator (e.g., between Buenos Aires and Mexico City) have been sustained on the 6-m band (near 50 Mc) during evening hours of equinoctial periods at times of sunspot maximum, suggesting unusual ionospheric conditions in the equatorial zone.

To provide a systematic study of ionospheric irregularities and other effects which might support ionospheric propagation at 50 Mc across the magnetic equator, an IGY experiment was set up in South America with circuits as illustrated in figures 10 and 11. A basic E-scatter circuit was installed between Arequipa and Trujillo, Peru, while the long-distance circuit between Antofagasta, Chile and Guayaquil, Ecuador was operated to observe scattering associated with effects in the lower F region. These circuits centered on the magnetic equator at Huancayo, recorded over a 1-yr period so as to enable the study of seasonal and diurnal variations as well as more rapid changes in the equatorial ionosphere. R. Cohen and K. L. Bowles, project leaders, have reported earlier [10], in some detail, the circuit arrangement and the objectives of the experiment, as well as some results of early data analysis [11]. The core of this experiment, the transmitting and receiving stations cited, was comprised of four field laboratory trailers, while the equipment installed at the cooperating laboratories in Huancayo, Clorinda, Argentina, and Sao Paulo, Brazil was housed in existing facilities. The west-coast observations began in November 1957, while the transcontinental ones were underway in April 1958, and all operations continued until 1 December 1958. The basic equipment for the forward scatter experiments consists of large rhombic antennas, high-powered (3 kw or 20 kw interchangeable) transmitters, and sensitive narrow-band receivers. The transmitting stations at Antofagasta, Chile (49.960 Mc), and Arequipa, Peru (49.920 Mc) and the receiving stations at Huancayo, Peru, Trujillo, Peru, and Guayaquil, Ecuador, employed 1000-ft rhombic antennas directed along the several great-circle paths. One transmitting station, Huancayo, Peru (49.880 Mc), and the two receiving stations at Clorinda, Argentina, and Sao Paulo, Brazil utilized stacked pairs of 5-element Yagi antennas. The Huancayo installation also had a small low-powered 100-w transmitter feeding a Yagi antenna directed toward Guayaquil. This circuit is intended for detection of the sporadic E between Huancayo

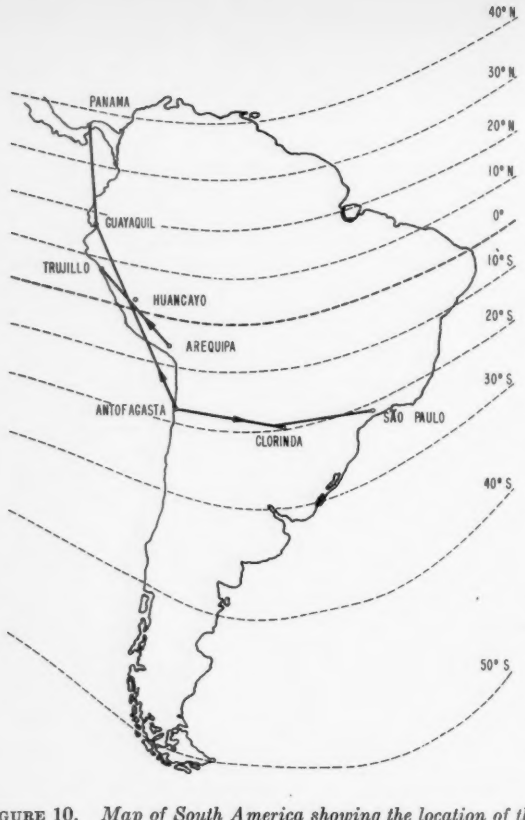


FIGURE 10. *Map of South America showing the location of the equatorial forward scatter stations.*

and Guayaquil. By using 100-cps bandwidth IF filters in the Guayaquil receiver, it proved possible there to record signals from Huancayo nearly 100 percent of the time, in spite of the low power. All receivers used a 12-sec integration in the automatic-gain-control circuit as a means of reducing the extreme variations on the chart due to normal fading. The quantity recorded was the AGC voltage, which is a good statistical estimator of the running median value of received signal power.

Two east-west circuits were installed along the 20° isocline, to receive signals from Antofagasta at Clorinda, Argentina via the *E* region and at Sao Paulo, Brazil via the *F* region. These circuits were installed to compare equatorial east-west propagation, insofar as possible, with the north-south propagation, as well as to check on the southern range of latitude over which equatorial effects are encountered. The presence of transmitters and receivers appropriately spaced permitted the reception of signals at Huancayo from Antofagasta, at Trujillo from Antofagasta, at Guayaquil from Arequipa in addition to the basic circuits previously mentioned and the Huancayo-Guayaquil circuit. Thus there are three symmetrical *E*-scatter paths,

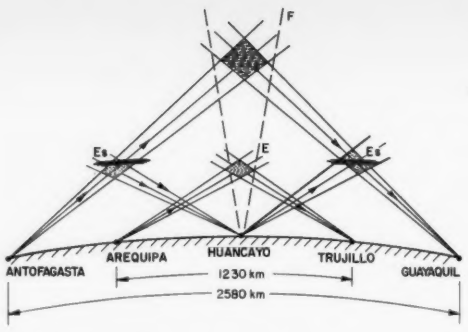


FIGURE 11. *Cross section of the equatorial forward scatter circuits in South America.*

as well as two asymmetrical paths, which permit the simultaneous observation of *E*-region effects at various latitudes. Ionospheric effects in this extensive region near the magnetic equator are particularly sensitive to latitude. Further details concerning the same ionospheric regions are provided by the records from the close-spaced chain of ionospheric sounders along the west coast of South America.

A preliminary comparison of the records obtained over the six circuits along the west coast of South America shows a number of interesting results; however, the complete picture will only be understood after a thorough analysis has been made. Over the two paths, Antofagasta to Guayquil and Arequipa to Trujillo, the short-term daytime variations in the signal intensity appear to correlate very closely. Indeed, over all six circuits, the daytime signal traces appear to correlate reasonably well while nighttime correlations are poor. Comparisons with the vertical incidence soundings at Huancayo suggest that these daytime signals result from scattering by equatorial sporadic *E*. A close correlation exists, during certain periods studied, between the signal over the Arequipa to Trujillo path and the *H* component of the Huancayo magnetogram. This is particularly the case when the signal correlation over all six paths is good. Since the *H* component has been interpreted as a measure of the equatorial electrojet, a close relationship is now indicated between VHF signals and the electrojet, and perhaps also between equatorial sporadic *E* and the electrojet. The Antofagasta to Huancayo and the Huancayo to Guayaquil records both show sporadic *E* signals similar to those over circuits in middle latitudes, while the curious aspect is that the circuit from Arequipa to Trujillo almost never exhibits this form of sporadic *E* propagation.

Often during the evening and nighttime hours on the Antofagasta to Guayaquil path (but not on the Antofagasta to Trujillo path) there is observed a strong signal increase which appears to correlate well with the occurrence of spread *F* on the vertical incidence sounder at Huancayo. While it had been expected that *F*-region scatter might take place over

this long-path circuit much of the time, this is evidence that it can occur at night.

The cooperation and participation of the IGY committees of Argentina, Brazil, Chile, Ecuador, and Peru was an indispensible part of this experiment, and the assistance by the government of these countries was greatly appreciated.

5. IGY Radio Noise Network

Radio noise is the basic limitation to radio reception, and in particular, the background of radio noise originating from natural causes in the atmosphere is often a severe restriction to good radio communication. Therefore, a detailed knowledge concerning the distribution of noise levels throughout the world and an understanding of the noise characteristics is fundamental to the advancement of communication techniques. For this reason, during the IGY, a worldwide network of stations recording radio noise levels was put into operation by NBS under the supervision of W. Q. Crichlow, R. T. Disney, and C. A. Samson.

In any radio noise measurements the following three parameters are useful: (1) The mean power, (2) the mean envelope voltage, and (3) the mean logarithm of the envelope voltage. The mean power averaged over several minutes is the basic parameter, and is expressed as an effective antenna-noise figure, defined as the noise power available from an equivalent lossless antenna in decibels above the thermal-noise power available from a passive resistance. Measurements of these parameters were made with the NBS radio-noise recorder, Model ARN–2, which has an effective bandwidth of 150 to 300 cps and uses a 21.75-ft vertical antenna and elevated ground-plane system. There are eight channels used between 13 kc and 20 Mc, two channels being recorded simultaneously. A 15-min recording is made on each frequency each hour, and is taken as representative of the noise condition throughout the hour. In addition to measurements made with the ARN–2 in the worldwide network instantaneous amplitude probability distributions were obtained at Boulder, Colo. These distributions provide information concerning the detailed character of the noise from which its interference properties can be determined.

Preliminary results would indicate that the worldwide predictions as presented by the CCIR report No. 65 are too simplified. The behavior of radio noise, as a function of the frequency, time of day and season, differs widely from location to location. Specific local conditions need to be well understood in order to make accurate predictions.

As a few examples of this one can point out the following. Afternoon noise levels in areas of high local thunderstorm activity may be 20 to 30 db higher than indicated by the predictions. Boulder, Colorado is a good example of this, since it is located near one of the most active thunderstorm regions in the United States—the front range of the Rocky Mountains. On certain days the variation may be as much as 70 db, on some frequencies. In contrast, at Kekaha in the Territory of Hawaii, the minimum noise levels for the low frequencies tend to occur near sunset instead of a few hours after sunrise as at most stations. As a result, the 160-kc afternoon noise levels appear to be 20 to 30 db lower than predicted at this location. The nighttime noise levels at Boulder are often lower than predicted while at Kekaha they agree rather well with the predictions.

The data recorded during the IGY at the stations listed in table 2 have been published in IGY data booklets and quarterly summaries have been issued as NBS reports giving trends and tentative conclusions.

TABLE 2. *Radio noise stations*

Stations	Geographic coordinates	
	Latitude	*Longitude*
Balboa, C.Z.	9.0°N	79.5°W
Bill, Wyo.	43.2 N	105.2 W
Boulder, Colo.	40.1 N	105.1 W
Byrd Station, Antarctica	80.0 S	120.0 W
Cook, Australia	30.6 S	130.4 E
Enköping, Sweden	59.5 N	17.3 E
Front Royal, Va.	38.8 N	78.2 W
Ibadan, Nigeria	7.4 N	3.9 E
Kekaha, Kauai, T.H.	22.0 N	159.7 W
Ohira, Japan	35.6 N	140.4 E
Pretoria, U. of S. Africa	25.8 S	28.3 E
Rabat, Morocco	33.9 N	6.8 W
São Jose dos Campos, Brazil	23.3 S	45.8 W
Singapore, Malaya	1.3 N	103.8 E
Thule, Greenland	76.6 N	68.7 W

6. Radio Satellite Observations at NBS

Within 12 hr of receiving the announcement on 4 October 1957 concerning the launching of Sputnik I, two groups at the Central Radio Propagation Laboratory, under the direction of R. S. Kirby, J. M. Watts, and R. S. Lawrence, had modified certain existing pieces of equipment to receive the 20- and 40-Mc signals from the first artificial earth satellite. Soon thereafter, a third group, under the direction of J. W. Herbstreit and M. C. Thompson, Jr.[12] made measurements using a continuous phase-difference measuring technique developed earlier for studying atmospheric turbulence in the troposphere. The instrumentation consisted of two similar Hamerlund SP–600 receivers operating on 20 (or 40) Mc. Continuously recording phase meters were used which compared the af tones from the two receivers. At first, half-wave dipoles were used, but later circularly polarized antennas were employed. Two effects were contained in the records: regular geometrical changes such as those related to the orbit and irregular changes presumably related to ionospheric turbulence and refraction. The group under R. C. Kirby and J. M. Watts used dipole antennas to measure the field intensity on 20 and 40 Mc from which it was estimated that the total power emitted by Sputnik I was of the order of 1 w. This value was later confirmed by the Soviet scientists. The group under the direction of R. S. Lawrence and E. R. Schiffmacher modified the receivers in the 40-ft dishes used for the radio star scintillation measurements in order to receive the 40-Mc satellite

signal. The two dishes are separated by a spacing of 1,500 ft. Within a few hours, dipole antennas were set up at twice the spacing of the dishes in order to receive the 20-Mc signal on an electrically equivalent baseline. The phase-switching and detecting equipment, that had previously been used for the Jupiter project, was put in operation here. Measurements were made of the direction of arrival of the 20 and 40 Mc signals. Simple ionospheric theory suggests that simultaneous measurements of refraction, at 20 and 40 Mc, should be in the ratio of 4 to 1. If this were correct, simultaneous measurements on the two frequencies could be used to deduce the true satellite direction. Observations have shown, that in fact, the irregular structure of the ionosphere invalidates this procedure.

Following the initial "crash" program more permanent interferometer arrangements were made at Table mesa near Boulder. Pairs of half-wave dipoles, with horizontal east-west polarization at one third wave length above the ground, were set up at the corners of equilateral triangles of 3,000-ft and 150-ft spacing. Observations were made at both 20 and 40 Mc. In addition circularly polarized antennas were installed. The output can be recorded on magnetic tape or on two 6-channel Sandborn recorders. The equipment is fully automatic and can be set for automatically coming on during the passage of a satellite beacon transmitter at these frequencies. Two Yagi antennas on a frequency of 108 Mc have also been installed and a vertically polarized antenna for 20 Mc. The satellite ionospheric measurements program at the National Bureau of Standards is now under the direction of C. G. Little. With this arrangement of antennas measurements of angle of arrival, Faraday rotation, polarization, scintillation, diversity effects, and Doppler frequency changes have been made.

Inspection of the Faraday fading records of the 20 Mc signals made during the first 4 months of satellite 1958 δ show the following systematic variations in the fading rate: (a) The fading rate is slowest when the satellite is in the north and greatest when it is in the south; (b) the greater height of the satellite during the north-to-south passes, and hence the longer ionospheric paths of the radio signals, accounts for the greater fading rates during these passes than during the south-to-north passes; (c) nighttime passes are characterized by a fading rate that is much slower than for daylight passes at the same height. When the fading rate was compared with the square of the *F*-layer critical frequency, taken from the Boulder ionograms, it shows that the daily variations in fading rate are largely due to the daily variations in the ionization content of the ionosphere. Many of the Faraday-rotation rates, rather than varying smoothly throughout a single passage of the satellite, vary irregularly and indicate the existence of ionospheric irregularities. In some instances these irregularities are of the order of 100 km in lateral extent. Calculations, using the 20 Mc signal from 1958 δ–2, show that the observed irregularities in the Faraday rotation rate can be accounted for if the integrated electron content, along the line of sight, per square centimeter column, varies irregularity with variations of the order of 10^9 during a 1-km change (horizontal) in satellite position. From a determination of the number of complete rotations of the plane of polarization of the radio signal, the total electron content, up to the height of the satellite, can be calculated. For a pass of satellite 1958 δ–2, at a height of about 236 km, low in the F region, the integrated electron content in a vertical column was estimated as 2.7×10^{12} electrons/cm^2 column.

In conjunction with the satellite observations a computer program is being set up for tracing the paths of the rays between the satellite and the receiver. The ordinary and extraordinary rays are traced separately through the ionosphere and the difference in their phase paths gives the instantaneous orientation of their polarization vector. Actual ionospheric profiles from ionograms are used up to the *F*-region peak, but various assumed profiles are tried for the upper ionosphere. The ray-tracing program is coupled to a program which gives successive positions of the satellite along an elliptical orbit. The combination of these two programs gives a complete picture of the Doppler frequency, the Faraday rotation, and the angle of arrival of the satellite signal. A comparison between the direction of arrival and line of sight to the satellite gives a measure of the angular errors introduced by the ionosphere.

7. Airglow Observations

The atoms and molecules of the earth's upper atmosphere continuously radiate a flux of light and radiation. During the daylight hours this flux is swamped by a much stronger energy flux from the sun, but at night the airglow emitted from the high atmosphere can be easily detected by sensitive photometers and on occasion it can be seen by the human eye. Closely related to the airglow emission, and perhaps a part of the same phenomenon, is the well known and often observed aurora associated with high-latitude belts on the earth. Because of its wide distribution and constant occurrence, it was logical that a worldwide study of the airglow be undertaken as a part of the IGY synoptic program. F. E. Roach has directed the NBS airglow program involving observation stations at Fritz Peak, Colorado and Rapid City, South Dakota. In addition, the group constructed photometers for College, Alaska, Huancayo, Peru, and Sidney, Australia. The photometer itself is of a hybrid design combining the Lyot birefringent filter principle, as adapted by J. Blamont and A. Kastler [13] and used by R. B. Dunn and E. Manring [14] with the telescope system designed by F. E. Ross. The telescope has a 6-in. objective, a focal length of 13 in., and a 4-degree field of view. A series of color filters permit the selection of one of the following emissions characteristic of the night airglow: OI ($^1S_0$–$^1D_2$) at 5577 A, OI ($^1D_2$–$^3P_1$) at 6300 A, and NaI ($^2P_{3/2}$–$^2S_{1/2}$, $^2P_{1/2}$–$^2S_{1/2}$) at 5890 and 5896 A. In addition there are many strong bands of emission

in the infrared from the OH molecule. The photometer is programmed to sweep the sky in successive almucanturs at angles of 80°, 75°, 70°, 60°, 40°, and the zenith. The photometer sweeps the entire sky in 4 min at each wavelength and returns once each 15-min interval to repeat the wavelength sequence. The data are recorded on a strip chart recorder and transferred to punch cards for processing on the IBM 650.

In addition to the photometer described above, a simple photometer of the Roach-Elvey type was installed at Thule, Greenland in October 1958 for special observations at the magnetic pole. There was also installed an all-sky camera and a patrol spectrograph at Fritz Peak and at Rapid City.

A number of notable results have come from this IGY airglow program at the Central Radio Propagation Laboratory. Since a number of publications have appeared describing these results they will only be referred to here and not described in detail.

An analysis of the nightglow observations at 5577-A shows the emission to come from a height of approximately 100 km [15]. Direct measurements of the altitude of the night airglow obtained from rockets [16] show the maximum of the 5577-A layer to be near 98 km. Systematic isophote airglow maps of the night sky throughout many nights have been prepared and a careful analysis of these maps made. This analysis shows the sky to contain discrete airglow cells of 5577-A emission, having a characteristic diameter of 2500 km and moving with translational speeds of about 100 m/sec. These airglow cells appear to exhibit a rotational period of about 5 hr.

In order to investigate the possibility that the 5577-A night airglow, present over the earth as a whole, may be closely related to the aurora, which is prevalant in high-latitude regions, the CRPL investigators have made a statistical analysis concerning the frequency of occurrence of these phenomena [17, 18]. Although this type of statistical treatment is open to some question it may give a clue as to whether or not these two phenomena, the airglow and the aurora, have a common origin.

8. World Data Center A—Subcenter for Airglow and Ionosphere

Located in the Boulder Laboratories of the National Bureau of Standards is the United States IGY World Data Center A, Subcenter for Airglow and Ionosphere Data. The Data Center is under the supervision of W. B. Chadwick and S. M. Ostrow. To this subcenter, all airglow and ionospheric stations in the western hemisphere and Australia and New Zealand send their data for copying and filing. In this manner, a central file of the results of the IGY observations will be made available for research and study for many years to come. It was the responsibility of this subcenter during the IGY to send or arrange to be sent copies of the data to World Data Centers B, C1, and C2 located in the USSR, England, and Japan respectively. These centers in return send data from the laboratories in their area of responsibility to Boulder. This subcenter, in addition to maintaining files of the data, prepares catalogues and indexes of the data, makes copies of the data available to scientists, organizations, and others requiring its use, and maintains at the Data Center office space where visiting research workers can come for study.

Although not all stations supply the same amount of data to the Data Center, the following items were contained in the recommendations of the First Report of the Special Committee on Worldwide Soundings:

1. Hourly values of f_0F_2, $h'F_2$, (M3000)F_2, f_0F_1, $h'F$, (M3000)F_1, f_0E, $h'E$, F-min, f_0Es, $fbEs$, $h'Es$, and other types of Es.
2. All median values of f_0F_2, (M3000)F_2, and f_0Es and other median values.
3. All stations were to produce f plots for Regular World Days, Special World Intervals, and any other occasions on which quarter-hourly or more frequent observations are scaled. High-latitude stations and a number of other key stations were to supply f plots for every day.
4. Ionograms at least for the Regular World Days and Special World Intervals.

In addition, other special types of data will eventually be on file in the Data Center, such as whistler data, radio-noise data, equatorial forward-scatter data, and data from other unique experiments. For the airglow studies zenith-intensity digital data for several stations, for the oxygen emissions at 5577 A and 6300 A, sodium at 5893 A, and the near infrared OH emission, is coming into the Data Center. The stations included in the American chain are Thule, Saskatoon, Rapid City, Fritz Peak in Colorado, Sacramento Peak in New Mexico, Tantazantla and Huancayo, Peru.

A steady flow of data has been coming in to the Data Center during the latter part of the IGY and it will be more than a year after the close of the IGY period before most of the data is in. The Data Center has received sizable amounts of data from the other World Data Centers. In addition to the IGY data, the Data Center has available in Boulder an extensive archive of data covering many years of records from stations throughout the world.

Workers desiring data may obtain it for the cost of reproduction from the World Data Center A, Subcenter for Airglow and Ionosphere, Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colorado.

The author wishes to express his appreciation to the project leaders and others for their assistance in the preparation of this article.

9. References

[1] R. W. Knecht and D. W. Schlitt, Early results from the equatorial close-space chain of ionospheric vertical soundings stations, Proceedings of the Vth CSAGI Assembly, (Moscow 1958), Annals of the IGY.

[2] K. G. Budden, A method for determining the variation of electron density with height (N(z) curves) from curves of equivalent height versus frequency (h′, f) curves, Rept. Cambridge Conf. Ionospheric Phys. p. 332 (Phys. Soc. London 1954).

[3] E. K. Smith, Worldwide occurrence of sporadic E, NBS Circ. 582, March (1957).

[4] S. Matsushita, Intense Es ionization near the magnetic equator, J. Geomagnetism and Geoelectricity **3,** 44 (1951).

[5] H. Odishaw, International Geophysical Year, Science **128,** 1599 (1958).

[6] S. Matsushita, Lunar effects of the equatorial Es, J. Atmospheric and Terrest. Phys. **10,** 163 (1957).

[7] R. W. Knecht, submitted for publication, Journal of Atmospheric and Terrest. Phys.

[8] R. W. Knecht, R. E. McDuffie, and Y. Aono, Preliminary results from U.S.-IGY antarctic network of ionospheric vertical soundings stations, Proc. of the Vth CSAGI Assembly, Moscow, Aug. (1958) Annals of the IGY.

[9] R. Bateman, J. W. Finney, E. K. Smith, L. H. Tveten and J. M. Watts, International Geophysical Year observations of F-layer scatter in the Far East, J. Geophys. Research **64,** (1959).

[10] K. L. Bowles, and R. Cohen, NBS equatorial region vhf scatter research program for the IGY, QST, **41,** 11–15 (August 1957).

[11] R. S. Cohen, and K. L. Bowles, 50-Mc oblique transmission experiment near the magnetic equator, Presented to URSI–IRE Spring Meeting, Washington, D.C. (1958).

[12] J. W. Herbstreit and M. C. Thompson, Jr., Continuous phase difference measurements of earth satellites, Proc. IRE **46,** 1535 (1958).

[13] J. Blamont and A. Kastler, Réalisation d'un photomètre photoélectrique pour l'étude de l'émission crépusculaire de la raie d du sodium dans la haute atmosphère, Ann. géophys. **7,** 73, (1951).

[14] R. B. Dunn and E. R. Manring, A recording night sky photometer of high spectral purity, J. Opt. Soc. Am. **45,** 899 (1955).

[15] F. E. Roach, L. R. Megill, M. H. Rees, and E. Marovich, The height of nightglow 5577, J. Atmospheric and Terrest. Phys. **12,** 171 (1958).

[16] R. Tousey, Rocket measurements of the night airglow, Ann. géophys. **14,** 186 (1958).

[17] F. E. Roach, J. W. McCaulley, and E. Marovich, Origin of [OI] 5577 in the airglow and the aurora, J. Research NBS **63D,** 15 (1959).

[18] F. E. Roach, J. W. McCaulley, and C. M. Purdy, Comparison of absolute intensities of [OI] 5577 auroral and subauroral zones, J. Research NBS **63D,** 19 (1959).

Boulder, Colo. (Paper 63D1–1).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Origin of [OI] 5577 in the Airglow and the Aurora

Franklin E. Roach, James W. McCaulley, and Edward Marovich

(January 30, 1959)

The distribution of 5577 zenith intensities at stations in the subauroral zone is found to be unimodal with no discontinuity at the visual threshold. This is interpreted as evidence that 5577 airglow and 5577 aurora may have a common origin.

1. Introduction

It is customary to refer to [OI] 5577 as airglow if it is invisible and as aurora if it is visible. Thus, the *physiological* visual threshold (at approximately 1,000 rayleighs [1] for extended objects) has been assumed to have a *physical* significance in the upper atmosphere with different mechanisms operative above and below the threshold. The purpose of this paper is to inquire into the existence of a physical criterion to distinguish between 5577 airglow and 5577 aurora.[2]

2. Analysis of St. Amand and Ashburn

One approach to the problem has been offered by St. Amand and Ashburn.[3] They suggest that a study of the statistical distribution of absolute 5577 intensities at a given location may yield information as to a common or separate origin for airglow and aurora. If there is a common origin, it might be expected that the distribution of intensities would be unimodal with a continuous gradation from the invisible (airglow) to the visible (aurora). If, on the other hand, the origins are separate and distinct, one would expect a discontinuity [4] of some kind in the distribution curve, the most likely being a bimodal distribution.[5] In figure 1, schematic representations of the two possibilities are shown. The two sketches on the left indicate a schematic bimodal distribution. In the subauroral zone, auroras are infrequent and this is indicated by the relatively smaller area under the aurora curve than under the airglow curve. In the auroral zone, the relative areas are shown reversed. Actually, the wings of the two distribution curves may overlap, but the essence of the argument is that there should be a discontinuity in the composite distribution if the two phenomena are independent in nature and origin. In the right-hand part of figure 1, schematic distributions are shown for the single-origin hypothesis, their critical characteristic being the lack of any discontinuity.

In figure 2, a replot of the histogram of St. Amand and Ashburn is shown with a regrouping of the data to correspond to class intervals of 50 R rather than the 10 R that they used. They put an approximately normal curve through the plot on the basis of which they were able to state that an intensity corresponding to a just visible aurora should occur only "once or twice in geologic time." Because auroras do occur occasionally, though rarely, in southern California, the implication is that an aurora is a completely separate phenomenon from the airglow (5577).

Two comments are in order on the analysis by St. Amand and Ashburn: (1) Since they used mean nightly values of 5577 intensity, their prediction must refer to cases in which the *mean* nightly intensity will be of auroral intensity, which is, of course, much less probable than that an aurora will occur for a short time during a night; (2) the assumption of a normal curve is a critical feature of their analysis. Even a slight positive skewness in the distribution would change the prediction radically. As is shown in the next section, positive skewness is a general characteristic of all the distributions for which a substantial amount of data is available.

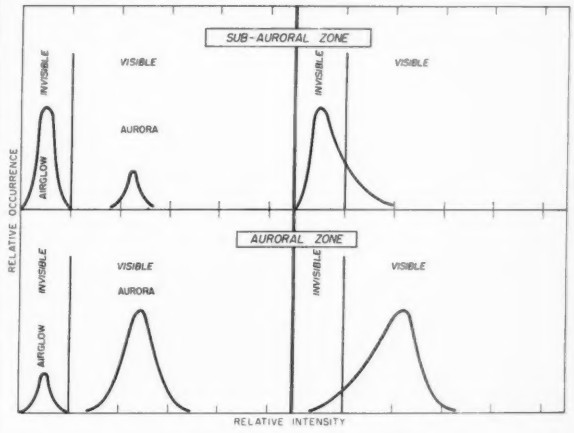


FIGURE 1. *Schematic representation of the common origin hypothesis (right) and the separate origin hypothesis (left) for the 5577 airglow and the 5577 aurora.*

[1] For a discussion of the "rayleigh" unit of intensity see D. M. Hunten, F. E. Roach, and J. W. Chamberlain, A photometric unit for the airglow and the aurora, J. Atmospheric and Terrest. Phys. **8**, 345 (1956). If the surface brightness (B) is measured in 10^6 quanta/cm²·sec·steradian, then the intensity in rayleighs (R) is $4\pi B$. For bright auroras it is convenient to use the kilorayleigh (kR) as a unit.

[2] Throughout this paper the discussion is strictly limited to the single emission 5577 and should not be generalized to any other radiation.

[3] P. St. Amand and E. V. Ashburn, The frequency distribution of the intensity of aurorae and the night airglow for 5577 [OI], J. Geophys. Research **60**, 112 (1955).

[4] In this paper the word "discontinuity" is used with the general implication of a "break" rather than in the strict mathematical sense.

[5] It should be mentioned that bimodality almost certainly indicates heterogeneity, but there may be heterogeneity without bimodality.

In their note on the problem, St. Amand and Ashburn examined the distribution of the nightly mean intensities for 86 nights at Cactus Peak, Calif., calling attention to the need for a larger statistical sample in order to give a more definitive solution. During the International Geophysical Year, a much larger sample is available for analysis. In the present study, some of this accumulated information is used to reexamine the problem.

3. Analysis of the IGY Data

At the present time, there are in the files of the Boulder World Data Center hourly zenith airglow intensities from 15 stations for approximately the first half of the International Geophysical Year. The distributions of 5577 zenith intensities *without a single exception have a pronounced positive skewness* so that, although the higher intensities near and exceeding the visual threshold are rare, they do not drop to zero occurrence. Furthermore, the distributions are in all cases smooth with no indication of any discontinuity. The strength of the above empirical fact can be appreciated when it is recalled that the calibrations and instrumentation are quite independent at the several stations.

The first treatment of the data from any given station should be the privilege of those in charge of the station so the details are not given here for airglow observing stations not under supervision of the authors. Since the distributions at Fritz Peak, Colo., and at Rapid City, S. Dak., are typical of those obtained at the other locations, the results are illustrated for these two stations plus the composite results for all the 15 IGY stations without individual identification.

In table 1 are shown the percentage occurrence by class interval and the cumulative percentage occurrence for: (a) Cactus Peak for the mean nightly intensities according to St. Amand and Ashburn (see section 2); (b) Fritz Peak hourly zenith intensities from July 1957 to September 1958; (c) Rapid City hourly zenith intensities from September 1957 to July 1958; and (d) the composite results for 7515 individual zenith intensities from 15 IGY stations.

In figure 3, the distributions for Rapid City, Fritz Peak, and the composite for all IGY stations are shown. In all three cases, the departure from a symmetrical normal distribution is clear. The large sample (7515) for the composite IGY results gives an unusually smooth distribution with the positive skewness clearly evident.

4. Discussion

To summarize the results obtained, it is noted: (a) For the stations studied there is evidence for a unimodal distribution of intensities with no apparent discontinuities; (b) the distribution shows a positive skewness so that there is a definite smooth occurrence of intensities from the more probable airglow values near 200 to 300 R to and past the visual threshold near 1,000 R.

Referring now to the introductory remarks, the facts are noted to be consistent with the hypothesis that 5577 airglow and aurora have a common nature and origin, at least for the stations in the subauroral zone. A further numerical test of the hypothesis is whether the distributions predict the occurrence of visual auroras. According to E. H. Vestine [6] there

[6] E. H. Vestine, The geographic incidence of aurora and magnetic disturbance northern hemisphere, Terrestrial Magnetism and Atmospheric Elec. **49**, 77 (1944).

TABLE 1. *Intensity frequency distributions*

Range	Cactus Peak (nightly means)		Fritz Peak (hourly values)		Rapid City (hourly values)		IGY (hourly values)	
	Size of sample							
	86		832		593		7515	
	Class	Cumulative	Class	Cumulative	Class	Cumulative	Class	Cumulative
R	%	%	%	%	%	%	%	%
0 to 49	0.0	0.0	0.1	0.1	0.2	0.2	0.1	0.1
50 to 99	2.3	2.3	.2	.3	3.2	3.4	2.8	3.0
100 to 149	8.1	10.5	1.3	1.6	10.8	14.2	7.5	10.4
150 to 199	18.6	29.1	7.0	8.6	15.7	29.8	11.1	21.6
200 to 249	23.3	52.3	10.7	19.3	19.1	49.9	12.8	34.4
250 to 299	16.3	68.6	11.7	31.0	14.2	63.1	13.3	47.7
300 to 349	16.3	84.9	12.9	43.9	11.3	74.4	12.5	60.2
350 to 399	9.3	94.2	11.3	55.2	7.8	82.1	9.8	70.0
400 to 449	2.3	96.5	10.1	65.3	6.2	88.4	7.3	77.2
450 to 499	2.3	98.8	8.7	74.0	3.0	91.4	6.2	83.4
500 to 549	1.2	100.0	6.6	80.6	3.2	94.6	4.9	88.3
550 to 599			4.6	85.2	0.5	95.1	3.4	91.7
600 to 649			4.4	89.6	1.2	96.3	2.3	94.0
650 to 699			2.4	92.0	0.8	97.1	1.8	95.8
700 to 749			1.9	93.9	.7	97.8	1.4	97.2
750 to 799			1.2	95.1	.0	97.8	0.8	97.9
800 to 849			0.8	95.9	.7	98.5	.6	98.5
850 to 899			.8	96.7	.0	98.5	.4	98.9
900 to 949			.5	97.2	.3	98.8	.3	99.2
950 to 999			.4	97.6	.0	98.8	.2	99.4

is a 4.6 percent occurrence of auroras at Rapid City (Φ=53° N) and 2.4 percent for Fritz Peak (Φ=49° N).

It is not possible to make a definitive comparison between the predictions of Vestine and the present observations because the present intensities refer to the zenith and the curve of Vestine predicts an aurora in any part of the sky. However, it is probably significant that the cumulative frequencies leave a residual (see table 1) of 1.2 percent for Rapid City and 2.4 percent for Fritz Peak for intensities greater than 1,000 R, which are similar to the quantities from Vestine's curve. During the period covered, there were visible auroras at both stations, the observations of which are included in the general statistical results. That the residual percentage over 1,000 R for Fritz Peak is greater than that for Rapid City is surprising, but it is suggested that this may be due to the smallness of the samples. However, the maximum zenith intensities for Rapid City (5,870 R) and for Fritz Peak (2,830 R) reflect the greater intensity of individual auroras at Rapid City, as is to be expected because it is closer to the auroral zone.

5. Conclusion

Evidence has been presented consistent with the hypothesis that 5577 airglow and 5577 aurora are of a common origin. It is of critical importance as a further test of the hypothesis that similar analyses be carried out at stations closer to the auroral zone where visual auroras are more frequent. In the following paper, some results are shown from Thule, Greenland, near the center of the auroral zone.

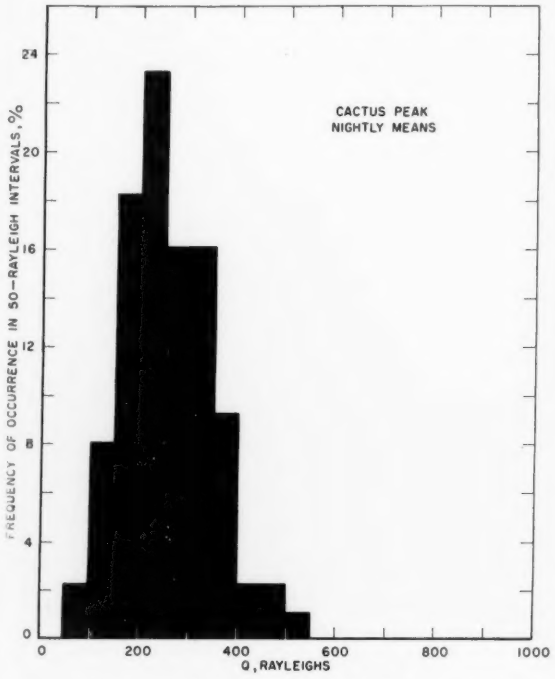


FIGURE 2. *Distribution of 5577 zenith intensities at Cactus Peak, Calif., for nightly means.*
From data of St. Amand and Ashburn, see footnote 3.).

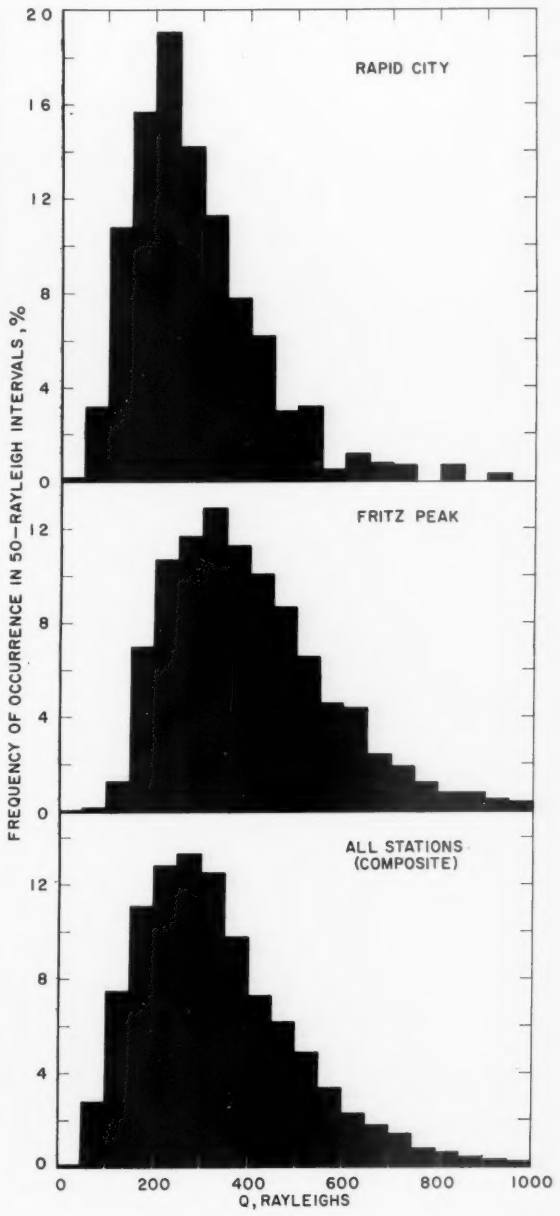


FIGURE 3. *Distribution of 5577 zenith intensities at Rapid City, S. Dak., and Fritz Peak, Colo., and for the composite of 15 IGY stations in the subauroral zone.*

The principal thesis of this paper is that there is no apparent discontinuity in 5577 distributions near the human visual threshold and that, therefore, this threshold should not be used to distinguish between two physical phenomena such as the airglow and the aurora. One often hears the statement from persons living close to the auroral zone that "there is some kind of an aurora every clear night" which can be paralleled by a similar statement by inhabitants of the subauroral zone that "there is 5577 airglow every clear night". Could it be that these statements correspond to physically similar situations at different average intensity levels? There remains the possibility of a break in the distribution *somewhere* in the four decades of 5577 intensity between a weak airglow (100 R) and an aurora IV (1,000 kR) but it does not show up at the visual threshold (1,000 R) in the subauroral zone. This paper is, therefore, a challenge to auroral observers to search for some quantitative factor in the aurora itself that can be used as a true discriminent.

BOULDER, COLO. (Paper 63D1–2).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Comparison of Absolute Intensities of [OI] 5577 in the Auroral and Subauroral Zones

F. E. Roach, J. W. McCaulley, and C. M. Purdy

(January 30, 1959)

The distribution of 5577 zenith intensities is compared for Fritz Peak, Colo., in the subauroral zone and Thule, Greenland, near the geomagnetic pole. The absolute intensity is bright enough to permit visual detection at the Colorado station during 2 percent of the time and at the Greenland station, 27 percent of the time. The distribution curves have a general similarity, suggesting a phenomenological similarity in the excitation mechanisms at the two stations.

1. Introduction

In the preceding paper[1] a discussion of [OI] 5577 zenith intensities in the subauroral region was given. In this contribution, some preliminary results are presented of measurements on the 5577 zenith intensity near the center of the auroral zone.

In October 1958, a photometer for the measurement of zenith intensities of 5577 was installed at Thule, Greenland, near the geomagnetic pole (geomagnetic latitude, $\Phi,=88°$). Two telescopes were used, one with an interference filter centered on 5577 and the other on 5300 A for control. The instrument had been carefully calibrated against the airglow photometers at Fritz Peak (geomagnetic latitude 49°) before installation in Greenland so that it is possible to make a direct comparison of the absolute intensities at the two stations.

2. Observational Data

The hourly values of 5577 zenith intensities at Fritz Peak were chosen for 125 nights from July 1957 to September 1958, a total of 832 readings. During the first two months of operation at Thule, there were 16 nights of observation. In general, the intensities changed more rapidly than at Fritz Peak, and measurements of the records were made for each ¼ hr, a total of 433 readings.

In table 1 are shown for the two stations, the direct and cumulative distributions of intensities according to 50-rayleigh class intervals up to 1,000 rayleighs (R).

In figure 1, the percent occurrence versus intensity is shown for both stations. In figure 2, the same data are shown with a logarithmic abscissa. In table 3, some of the statistical results are summarized.

3. Discussion

The following empirical facts are noted: (1) The occurrence (27%) of intensities greater than the visual threshold (≈1,000 R) is in substantial agreement with the results of visual auroral observations (20%) according to Vestine[2]; (2) there is no gap between the intensities conventionally referred to as airglow ($Q<1{,}000$ R) and those of an Aurora I (1,000 R$<Q<$10,000 R); and (3) the distribution curves for the two observing stations are similar in

TABLE 1. *Distribution of zenith intensities to the visual threshold*

Class interval	Fritz Peak		Thule	
	By interval	Cumulative	By interval	Cumulative
R	%	%	%	%
0 to 49	0.1	0.1		
50 to 99	.2	.3		
100 to 149	1.3	1.6		
150 to 199	7.0	8.6		
200 to 249	10.7	19.3		
250 to 299	11.7	31.0		
300 to 349	12.9	43.9		
350 to 399	11.3	55.2	1.2	1.2
400 to 449	10.1	65.3	3.0	4.2
450 to 499	8.7	74.0	5.5	9.7
500 to 549	6.6	80.6	9.7	19.4
550 to 599	4.6	85.2	6.5	25.9
600 to 649	4.4	89.6	8.5	34.4
650 to 699	2.4	92.0	9.9	44.3
700 to 749	1.9	93.9	8.3	52.6
750 to 799	1.2	95.1	7.4	60.0
800 to 849	0.8	95.9	3.7	63.7
850 to 899	.8	96.7	3.9	67.6
900 to 949	.5	97.2	2.1	69.7
950 to 999	.4	97.6	3.0	72.7
Above 999	2.4	100.0	27.3	100.0

TABLE 2. *Distribution of zenith intensities brighter than the visual threshold (Thule)*

Class interval	Percent	Class interval	Percent
R		*R*	
1,000 to 1,049	3.2	1,500 to 1,549	0.7
1,050 to 1,099	1.8	1,550 to 1,599	.7
1,100 to 1,149	3.2	1,600 to 1,649	.9
1,150 to 1,199	1.8	1,650 to 1,699	.2
1,200 to 1,249	1.4	1,700 to 1,749	.0
1,250 to 1,299	1.6	1,750 to 1,799	1.2
1,300 to 1,349	2.3	1,800 to 1,849	1.2
1,350 to 1,399	1.4	1,850 to 1,899	2.1
1,400 to 1,449	0.9	1,900 to 1,949	0.2
1,450 to 1,499	.9	1,950 to 1,999	.2
		Above 1,999	1.2

TABLE 3. *Summary of statistical results*

	Fritz Peak	Thule
Modes	325	650 (550, 675)
Percent brighter than 1,000 R	2.4	27.3
Percent of visible auroras (Vestine)	2	20
Percent fainter than 350 R	44	0
Percent between 350 and 1,000 R	54	73

[1] F. E. Roach, J. W. McCaulley, and E. Marovich, The origin of [OI] 5577 in the airglow and the aurora, J. Research NBS **63D**, (1959).

[2] E. H. Vestine, The geographic incidence of aurora and magnetic disturbance, northern hemisphere, Terrestrial Magnetism and Atmospheric Elec. **49**, 77 (1944).

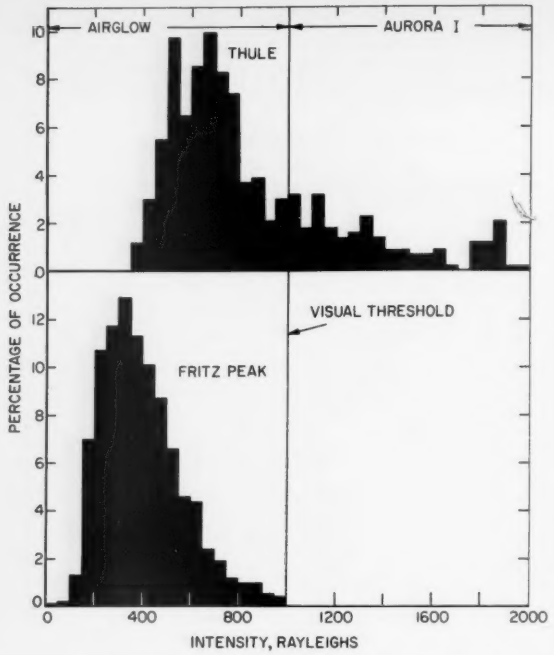


FIGURE 1. *Histograms showing distribution of 5577 zenith intensities (linear scale) at Thule and Fritz Peak.*

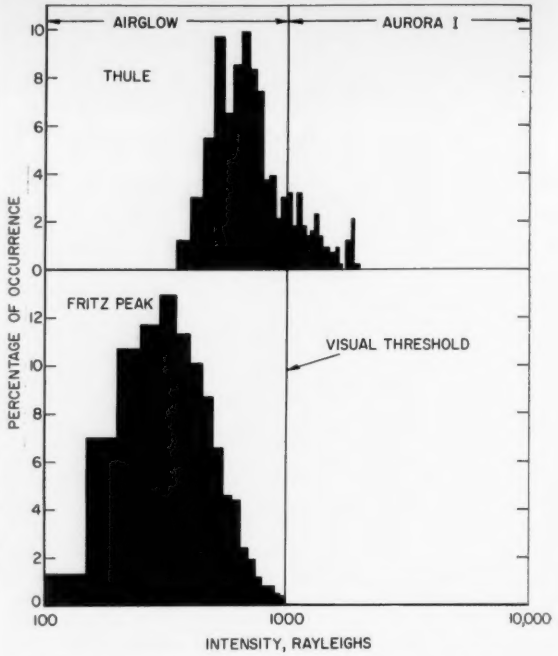


FIGURE 2. *Histograms showing distribution of 5577 zenith intensities (logarithmic scale) at Thule and Fritz Peak.*

appearance, roughly, but not exactly, as though the intensities at Thule were twice those at Fritz Peak.

Is there any evidence for a phenomenological distinction between the 5577 airglow ($0<Q<1{,}000$ R) and the 5577 Aurora I ($1{,}000$ R$<Q<10{,}000$ R)? The question is pertinent because of the fact that the visual threshold ($Q\approx1{,}000$ R in the green) is the distinction by definition between the 5577 airglow and an aurora with International Brightness Coefficient I. A widespread assumption is inherent in the current literature that the two phenomena are separate and distinct with different mechanisms responsible for the airglow (photochemical reactions) and the aurora (solar-terrestrial streams).

The evidence of this paper does not support the conjecture that the two phenomena are distinct but, rather, strongly suggests that a common physical mechanism should be sought for both.

The data so far accumulated do not give any information on the brighter Auroras II ($10{,}000$ R$<Q<100{,}000$ R$=100$ kR), III (100 kR$<Q<1{,}000$ kR), and IV ($Q>1{,}000$ kR), since none was observed of these intensities during the period of study. Near the auroral zone where auroras are reported nightly, it should be possible to determine (1) whether there is any evidence for a statistical distinction between Auroras I, II, III, or IV, or (2) whether for each station there is a distribution similar to those found for Fritz Peak and Thule. Probably the only published paper that bears on the question is that of Currie [3] who made visual observations at Saskatoon ($\Phi=60°.5$) with a Hulburt [4] photometer. According to Vestine's study, the auroral occurrence should be almost the same at Saskatoon as at Thule. It is of interest to note (table 4) that the zenith intensity observed by Currie was also on the average about twice that of subauroral regions, in agreement with present results from Thule and Fritz Peak.

TABLE 4. *Comparison of visual observations at Saskatoon and midlatitudes*

Location	Observed zenith intensity	Corrected for astronomical light	
	$m\mu L$	$m\mu L$	R
Midlatitudes (Hulburt [a])	130	68	309
Saskatoon (Currie [b])			
Late summer	185	123	560
Autumn	254	192	873
Winter	221	159	724
Quiet arc in NE (August)	198	132	600
Quiet arc in NE (later in year)	1,210	1,148	5,220

[a] See footnote 4. [b] See footnote 3.

The authors are deeply indebted to the Air Force Cambridge Research Center (Bedford, Mass.) personnel of which installed and operated the photometer at the Thule station. The personnel actively engaged in the operation were Maj. John B. MacWherter, 2d Lt. Don Smart, M. Sgt. John W. Hubin, and A2c Frank Pavlica, Jr.

[3] B. W. Currie, Night sky brightness at latitude 52°, Trans. Am. Geophys. Union **31**, 539 (1950).

[4] E. O. Hulburt, Survey of the brightness of the night sky, NRL Report N-3263 (1948).

BOULDER, COLO. (Paper 63D1–3).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Origin of "Very-Low-Frequency Emissions" [1]

R. M. Gallet and R. A. Helliwell[2]

(January 16, 1959)

Selective traveling-wave amplification in the outer ionosphere is postulated to explain very-low-frequency emissions, a class of very low-frequency (1 to 30 kilocycles per second) natural noise. By analogy with the mechanism of traveling wave tubes, low-level ambient noise in the outer ionosphere is amplified in streams of incoming ionized solar particles at frequencies for which the stream and wave velocities are equal. Required velocities are in the range 0.01 to 0.1 c (where c is the velocity of light). Streams with densities of the order of one electron per cubic centimeter would provide sufficient energy. Phenomena which can be explained qualitatively by the theory are the hiss, quasi-constant tones, dawn chorus and related transients, and very long trains of whistler echoes. A quantitative example shows how the theory can reproduce the general form of certain characteristic discrete spectra "hooks" of emissions, and how this leads to definite values of particle velocity and a law for the distribution of electron density in the outer ionosphere.

1. Introduction

When a high-gain audio amplifier is connected to a loop or long-wire antenna, unusual sounds of natural origin are often heard. The more impulsive sounds called "atmospherics" are produced by lightning and reach the receiver by traveling between the ground and the lower edge of the ionosphere. They are described by various terms such as "clicks," "tweeks," and "crashes."

Some of the energy from a lightning discharge also enters the ionosphere and is guided by the lines of force of the earth's magnetic field. As the energy propagates from one hemisphere to the other, it is dispersed and arrives at the receiver as a whistling tone, usually of descending frequency. Such signals are called "whistlers," and their theory has been developed in some detail (Storey [1],[3] Helliwell et al. [2], Ellis [3], and Maeda [4]).

There is yet another class of audiofrequency signals which we will call "very-low-frequency emissions". These appear to be related to magnetic disturbance and occur most frequently near the auroral zones. (Storey [1], Pope [5], Allcock [6], Dinger [7], and Watts [8].) One of the most common is the "hiss", a band of noise several kilocycles wide which is usually heard below 10 kc; it is occasionally accompained by one or more discrete tones in the same range of frequencies. Another is the so-called "dawn chorus", a combination of short (0.1–0.2 sec) rising whistles and warbling tones. Various isolated events, including rising and falling tones and combinations thereof can be observed. Very-low-frequency emissions are frequently associated with strong whistlers. Sometimes a whistler appears to "trigger" short periods of dawn chorus or hiss. Whistlers are sometimes preceded or followed by one or more rising tones.

It is the purpose of this paper to outline a theory of the origin of vlf emissions, based on selective traveling-wave amplification taking place in the outer ionosphere. The input signal is assumed to be provided by thermal radiation, whistler energy, or possibly Cerenkov radiation (Ellis [9]). It is assumed that the energy for amplification of the electromagnetic wave is provided by the streams of ionized solar particles (both positive and negative) which are thought to cause the auroras. These streams are assumed to penetrate the ambient ionization of the outer ionosphere and to be guided along the lines of force of the earth's magnetic field. The mechanism of amplification is essentially similar to that which takes place in an ordinary traveling wave tube (Pierce [10]), except that the "slow-wave" circuit (provided by the helix in the tube) is the ambient ionization of the outer ionosphere in the presence of the earth's magnetic field. Such a medium is dispersive, i.e., the wave velocity is a function of frequency, and at the very low frequencies considered here the velocity may be reduced as much as two orders of magnitude below that in free space. Furthermore, as shown by Storey [1] in connection with the theory of whistlers the energy of a vlf wave tends to be guided along the earth's magnetic field lines rather closely, and thus may follow the same path as the postulated particle stream. The wave normal, on the other hand, may make any angle with the field; and since the wave polarization is approximately circular, there will normally be a substantial component of the electric field of the wave which is parallel to the magnetic field. This component will act on the moving particle stream to produce the required sinusoidal modulation of charge density. For given values of the ambient plasma and gyrofrequencies, the stream velocity will approximate the phase velocity at two particular frequencies, and traveling wave amplification can take place. As the parameters of the medium change with position along the stream, the frequencies of amplification will change. This paper develops the conditions for interaction assuming the validity of the analogy with the traveling wave tube mechanism.

[1] The basic ideas described in this paper were presented to the Symposium on the Propagation of VLF Radio Waves, January 23–25, 1957, Boulder, Colo.
[2] Dr. Helliwell is affiliated with the Radio Propagation Laboratory, Stanford University, California.
[3] Figures in brackets indicate the literature references at the end of this paper.

2. Interaction Conditions

By analogy with the small-signal theory of traveling wave tubes (Pierce [10]), the necessary condition for coupling between the electromagnetic wave and the stream of charge is

$$v_p = v_s, \tag{1}$$

where v_p=phase velocity of wave, v_s=velocity of stream.

We must now deduce the conditions under which (1) will be satisfied. The phase velocity is cn, where c is the velocity of light and n is the refractive index in the medium. Since we need a phase velocity less than the velocity of light, we shall be interested in the "whistler-mode" of propagation (Storey [1]). For simplicity we shall consider the case in which the wave normal direction is approximately the same as that of the earth's magnetic field. The refractive index for the propagating mode is then given by

$$n^2 = 1 + f_p^2/\{f(f_H - f)\},$$

and the phase velocity is then simply

$$v_p = c\left[\frac{f(f_H - f)}{f(f_H - f) + f_p^2}\right]^{1/2}, \tag{2}$$

where f=wave frequency; f_H=gyrofrequency; $f_p = (81N)^{1/2}$=plasma frequency; with N=number of electrons per cubic centimeter when f_p is in kilocycles per second.

It is assumed that $f_p \gg f, f_H \gg f$. From (1) and (2), the frequencies of amplification are given by

$$f = \frac{f_H}{2}\left\{1 \pm \left[1 - \frac{\left(\frac{2v_s}{c}\frac{f_p}{f_H}\right)^2}{1-(v_s/c)^2}\right]^{1/2}\right\}. \tag{3}$$

If we restrict the discussion to cases in which the stream velocity is less than about 0.1 c, then $(v_s/c)^2 \ll 1$, and (3) can be written

$$f = \frac{f_H}{2}\{1 \pm [1 - (2P/f_H)^2]^{1/2}\}, \tag{4}$$

where $P = f_p(v_s/c)$ and is conveniently measured in kilocycles per second.

For interaction to occur, the solution to (4) must be real, and hence the condition for interaction is

$$f_H \geq 2P. \tag{5}$$

The maximum value of the wave frequency f is f_H, since the maximum real value of the radical in (4) is unity. The conditions for interaction are shown in figure 1, which is a plot of wave frequency f as a function of gyrofrequency f_H for different values of the parameter P. All curves are asymptotic to the line $f = f_H$ (ignoring the dashed curve for the moment).

As an example of the application of these curves, consider a stream of velocity 3,000 kmps which enters the outer ionosphere along an auroral line of force (near geomagnetic latitude 67°). The ambient plasma frequency is assumed to be constant and equal to 200 kc, and thus the parameter P=3 kc. When the stream reaches the point where $f_H = 2P$, interaction can begin at the wave frequency $f_H/2$, or 3 kc. Beyond this point two frequencies can be amplified, and as f_H increases the lower one approaches zero and the upper one approaches f_H, following the curve P=3 in figure 1. The locations of these frequencies along the dipole line of force are plotted in figure 2.

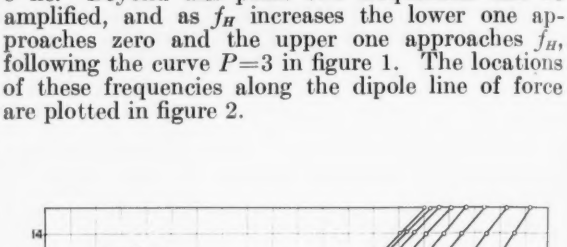


FIGURE 1. *Variation of excited wave frequency with gyrofrequency.*

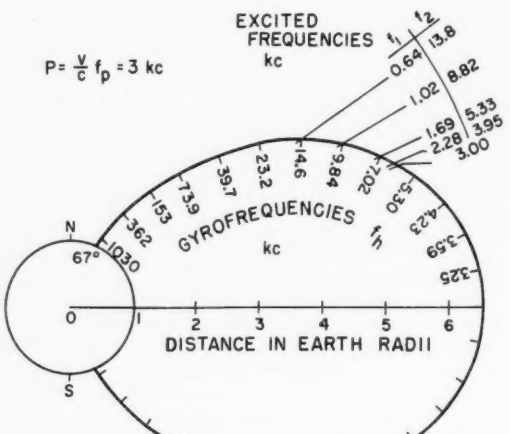


FIGURE 2. *Illustration of the theory of hiss, for the auroral line of force λ=67°.*

3. Stream Velocities

The stream velocity required for interaction is found from (4) and is given by

$$v_s = \frac{c}{f_p}\sqrt{f(f_H - f)}. \tag{6}$$

Values for the stream velocity for appropriate values of the medium parameters and the wave frequency are given in table 1. It is seen that the required stream velocities could vary roughly between 0.01 and 0.1 of the velocity of light. These velocities seem reasonable, since the lower value, 0.01 c, is comparable to the lower limit found by Meinel [11] in spectroscopic studies of hydrogen in the aurora. On the other hand, it appears necessary that electrons which penetrate to the E region must have velocities of approximately 0.3 c (Vegard [12]). Thus the velocities required for the traveling wave mechanism appear to lie in the same range as those ascribed to auroral particles.

4. Stream Energy Requirements

If traveling wave amplification is to be accepted as a possible mechanism, it is necessary that the power density of the incoming stream exceed that of the observed signals. A measurement of the strength of relatively strong hiss was made at Stanford, California (Stanford University [13]) on March 3, 1956, during a period of magnetic disturbance. The power density was of the order of 10^{-10} w/m². for a signal having a bandwidth of about 1 kc. Observations of whistlers indicate that the total attenuation from the source to the observer is not likely to exceed 20 db. Thus the power density of the signal as it leaves the amplifiying region is of the order of 10^{-8} w/m².

Let us now calculate the power density of the incoming stream. Since the particle densities and velocities of streams from the sun are not known, we must make a conservative estimate. From the work of Martyn [14] the stream is estimated to have a density of at least 1 electron/cm³ and a velocity of 3,000 km/s. The power density in such a stream is $(\frac{1}{2}mv^2)(Nv)$, where m is the particle mass, v the velocity, and N the number of particles per unit volume. For electrons, $m = 9.11 \times 10^{-31}$ kg, and the power density of the stream is thus about 10^{-5} w/m². The required efficiency of conversion from kinetic energy to electromagnetic energy is thus 0.1 percent, a low value compared with traveling wave tube efficiencies, which are of order 10 percent.

5. Comparison of Theory With Observations

Using the traveling wave hypothesis the hiss is explained in terms of a continuous stream of particles along which amplification occurs only for a restricted band of frequencies. The amount of gain at any one frequency will depend upon the interaction distance available, which in turn will depend on the way in which the plasma frequency varies with distance along the path. In a like manner the quasi-constant tones sometimes observed on different frequencies can be explained by assuming an ionization model in which the plasma frequency increases with gyrofrequency in such a way that f=constant over an appreciable length of path. Several different tones observed simultaneously could be explained by different streams flowing along lines of force of sightly different geomagnetic latitude, or possibly by streams of different velocities.

Transient phenomena, of which the dawn chorus is an example, are less easily explained. The traveling wave hypothesis can be adapted to the generation of transient signals by assuming that the incoming particles arrive in bunches. As the bunches travel along the lines of force, amplification takes place in the region of the bunch. However, the characteristic frequencies are amplified for a length of time which depends upon the length and velocity of the bunch. As the bunch moves, the frequency which is generated changes in accordance with changes in the parameters of the medium. The mechanism of formation of such bunches is not known; however, the visual phenomena observed during aurora suggest that solar streams may be bunched.

The interpretation of observed transient signals in terms of bunches of particles is complicated by the fact that the frequency-time relation of the received signal will depend on the integrated group delay from the point of origin to the receiver. The group delay, which is a function of frequency, clearly will affect the frequency-time dependence of the received signal. However, it cannot by itself explain the appearance of a given frequency at two different times, a phenomenon seen frequently in the dawn chorus. Such cases might better be explained in terms of variations in the ambient ioniza-

TABLE 1. *Calculated stream velocities for different parameters of amplifying region*

Gyrofrequency	Distance from surface	Minimum geomagnetic latitude	Plasma frequency	Wave frequency		
				2 kc	5 kc	10 kc
f_H= ,000 kc	Regular layers	Middle latitudes	4,000 kc	0.011 c	0.018 c	0.025 c
			1,000 kc	.044 c	.071 c	.1 c
f_H= 00 kc	~1.0 earth radii	44°	1,000 kc	.014 c	.022 c	.030 c
			200 kc	.070 c	.11 c	.15 c
f_H= 0 kc	3.6 earth radii	61°	200 kc	.02 c	.025 c	
f_H= kc	9.0 earth radii	65°	200 kc	.014 c		

tion density in the interaction region. For example, consider the hypothetical model, $f_p = 200 \text{ kc} + f_H^2/(1 \text{ kc})$, where f_H is measured in kilocycles per second. A short bunch of charge (say, with dimensions of about 100 km) traveling through such a region would excite the frequencies given by the dashed curve of figure 1. The energy would be limited to short groups produced at different times. The curve of excited frequency versus gyrofrequency transforms to a somewhat similar shape in the frequency-time plane. Some distortion of this curve occurs at the receiver because group delay is a function of both frequency and location of source. No completely closed trace of this type has yet been observed, but many of the recorded signals first fall and then rise with time, corresponding to the lower portion of the dashed curve. Thus there is at least a qualitative resemblance between the theory and the experimental data.

The steady hiss and the sharply defined transients represent extremes among vlf emission phenomena. In between lie many forms which show varying degrees of irregularity. Their character often changes markedly in the course of a few hours. It is suggested that the observed variety of form is simply the result of variations in the degree of bunching of the charge in the streams.

The traveling wave hypothesis offers a convenient explanation of very long trains of echoes of a whistler. The echo intensity decreases very slowly with order, and may even increase. It is difficult to explain this behavior in terms of the passive properties of the medium. Discrete echoes in a single train have been detected for as long as 4 min. These rather rare events are usually accompanied by magnetic disturbance and the dawn chorus or hiss. The whistler components which are most prolonged often lie in the 3- to 4-kc range where the hiss is known to predominate. On the basis of these facts it is suggested that the low apparent attenuation in very long whistler trains can be attributed to traveling wave amplification somewhere along the path. The whistler energy propagates along the same path followed by the incoming solar stream and is amplified in the regions where the whistler *phase* velocity approximates the stream velocity.

One of the most puzzling phenomena is the interaction between loud whistlers and vlf emissions. If during a period of dawn chorus and whistler activity, an ordinary loud whistler should occur, it is frequently accompanied by one or more rising whistles (Storey [1]). In attempting to explain this connection in terms of the traveling wave concept, we suggest that the loud whistler may be capable of accelerating an incoming stream of particles. For energy to be transferred from the whistler to the stream, the wave must accelerate the particles in its electrical field. Equation (2) shows that in a region where f_p does not increase too rapidly with f_H, the phase velocity of a fixed frequency wave will increase with f_H. For our purposes the rate of change of whistler frequency with time is sufficiently small that we can assume the accelerating signal to be of constant frequency. Thus as the whistler travels down the field line, particles of a given velocity are literally picked up by the wave at every point where a wave component of the same velocity exists. As the wave continues to propagate, its phase velocity increases and the particles riding on the front of the wave are accelerated, as in a linear accelerator. With the additional kinetic energy which the stream acquires from the signal it would be more efficient as a traveling wave amplifier and might conceivably produce audible vlf emissions.

6. Conclusions

The rather close similarity between conditions in the outer ionosphere and those in a traveling wave tube lead to the hypothesis that traveling wave amplification might take place at very low audiofrequencies. Even though the traveling wave hypothesis may not account successfully for all of the observed phenomena, it is probably an important factor to be considered in whistler propagation as well as in the generation of vlf emissions. Confirmation of the hypothesis must await a more detailed study of the traveling wave-gain problem and somewhat more detailed information regarding the distribution of ionization in the outer ionosphere. Such data presumably will come from whistler analysis. This theory, if correct, should provide a useful new tool for a study of the densities, velocities, and locations of streams of charged particles in the outer ionosphere.

7. Appendix: A Quantitative Comparison of the Theory With Observation

7.1. Example Chosen for Comparison With the Theory

The theory presented above has been developed in extensive numerical applications by one of us (R. M. G.) to see if the theory is able to account in detail for spectra of discrete vlf emissions, using a minimum of reasonable assumptions.

There is considerable information on the distribution of electron density up to the height of maximum density in the ionosphere, but very little is known about densities above this height. In particular, the law governing the distribution of electron density with height is not known for this region. In these investigations several models of the distribution above the main body of the ionosphere were considered. They were: Constant electron density; a power law r^{-n} where r is the distance from the center of the earth and n was taken from 1 through 6; and an exponential distribution e^{-kr} where k is a constant. The power law was found to give the best representation of certain distinguishing features of the vlf emissions and also to be consistent with information obtained from parallel studies of whistlers

(dispersion constants and analysis of nose whistlers). The same criteria seem to require the exponent $n \simeq 3$. Furthermore, a power law distribution with $n \simeq 3$ appears to be justified theoretically in other research in progress on the gravitational equilibrium of extended atmospheres, at large distances from the earth's surface. It is perhaps significant for the physics of the exosphere that $n \simeq 3$ means that the electron density is nearly proportional to the strength of the earth's magnetic field.

An example of a computed spectrum (frequency versus time at the earth's surface) of waves emitted by a discrete cloud of ionized particles moving with constant speed parallel to the lines of force of the earth's magnetic field is shown in the upper part of figure 3. This example was chosen to approximately match the observed spectrum shown in the lower part of the figure. The latter is a typical member of the "Hook" class of vlf emissions. This class is one of the best defined, and the reproducibility of the emissions is remarkably good in general shape. However, no two hooks, even though separated by only a few seconds in time, are exactly the same. Therefore, the exact matching of any individual tone is of no particular importance. In the example shown here, no great accuracy in matching was attempted; the same computations made with a slightly lower speed for the cloud of particles (say 9,000 km/sec) would have resulted in a better match.

It is more significant to realize that with reasonable numerical values for the parameters, the theory leads very naturally to calculated spectra very similar to those observed in regard to values of frequency, duration, and general shape.

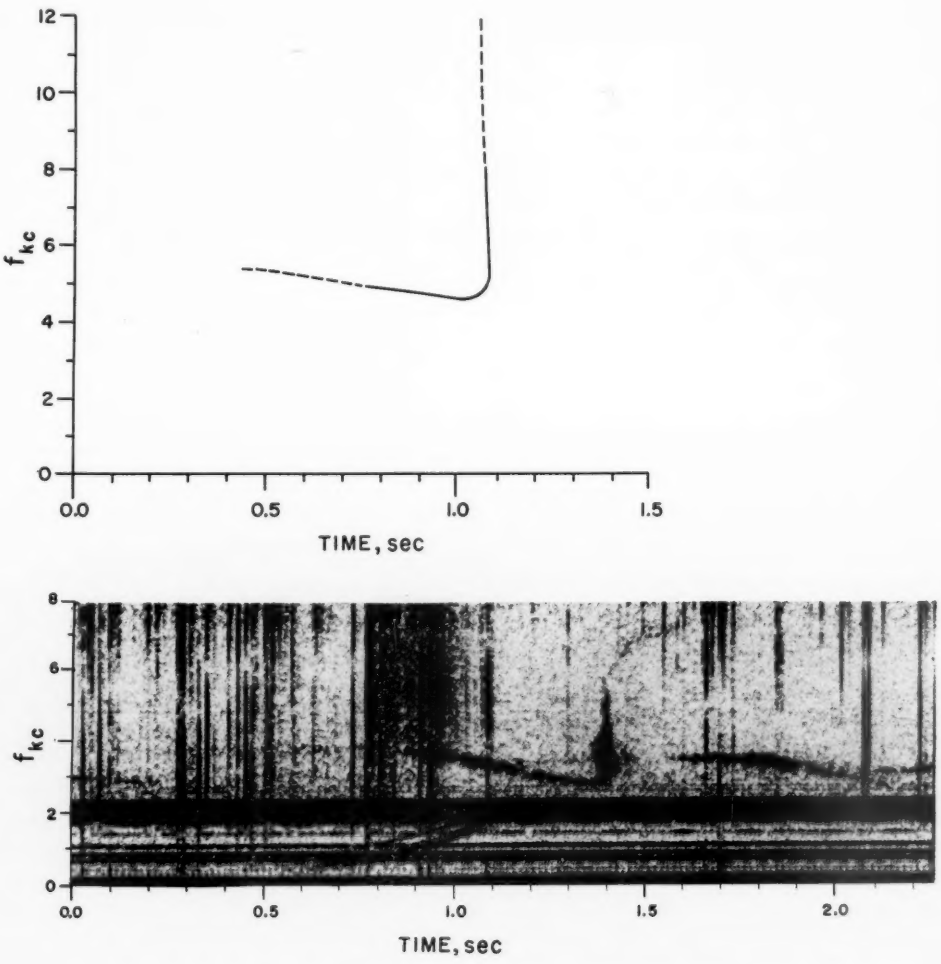


FIGURE 3. *Comparison of theory and observation (see appendix for details).*

7.2. Model of the Distribution of Electron Density

The model of electron density versus altitude used for the numerical calculations is illustrated in figure 4. Electron density (N_e/cm^3) is given on the right scale, and the corresponding local plasma frequency f_p in kilocycles can be read on the left scale.

In the exosphere, far above the earth's surface (above 1,000 km), rather than r^{-3}, the distribution chosen was:

$$Ne/H=\text{constant}, \quad \text{or} \quad f_p^2=Kf_H, \tag{7}$$

where H is the local magnetic field strength and f_H is the local electron gyrofrequency. Below 1,000 km, as a means of joining smoothly with a plausible ionospheric distribution, the above expression for f_p^2 is multiplied by the factor $\exp\{-(h-1{,}000)^2/\mathscr{H}_{ex}^2\}$, where h is the height in kilometers above the earth's surface. The scale height $\mathscr{H}_{ex}$ is taken as 200 km.

The ionospheric distribution is taken to be a Chapman-like layer

$$N_e=N_{max}\exp\{\tfrac{1}{2}(1-z-e^{-z}),$$

or

$$f_{p,\text{iono}}^2=f_{p,\max}^2\exp\{\tfrac{1}{2}(1-z-e^{-z})\}, \tag{8}$$

where $z=(h-h_{max})/\mathscr{H}$. The chosen value of N_{max} (10^6 cm^{-3}) corresponds closely to the critical frequency of the $F2$ layer at the place and time of the observation to be matched (Boulder, Colorado, October 26, 1956, 2230 UT or 1530 local standard time). The value of h_{max} was taken as 350 km. Above h_{max} the value of the scale height used was $\mathscr{H}_A$=200 km. Below h_{max}, $\mathscr{H}_B$=100 km was used.

Thus the total electron density for the model used is

$$f_p^2=f_{p,\text{iono}}^2+f_{p,\text{exo}}^2$$

$$=f_{p,\max}^2\exp\{\tfrac{1}{2}(1-z-e^{-z})\}$$

$$+\begin{cases} Kf_H\exp\{-(h-1{,}000)^2/\mathscr{H}_{ex}^2\} & \text{for } h<1{,}000 \text{ km} \\ Kf_H & \text{for } h>1{,}000 \text{ km}\end{cases} \tag{9}$$

Note that only a small range of values of the constant K is possible if the electronic density in the exosphere is not to be too large near the earth (smaller than the lowest values of N_{max}) nor too small to permit the propagation of whistlers at the greater distances. The value chosen here was 4,000 kc.

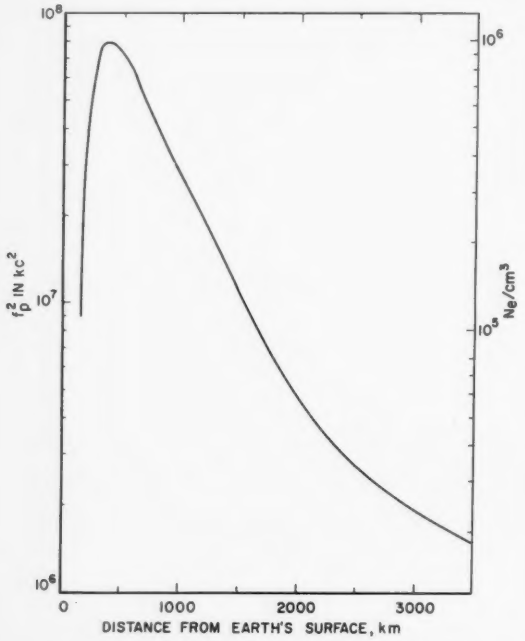


FIGURE 4. *Distribution of electron density used for the theoretical calculation shown in figure 3.*

7.3. Geometric Conditions

The calculations were made for an ionized cloud of particles moving along the line of force which intersects the earth's surface at the geomagnetic latitude λ=55°. The observations were made at Boulder at geomagnetic latitude 49°, but it is known from studies of whistlers that a signal propagating in the whistler mode can be received at places far from the foot of the nominal line of force along which the wave is propagated. For a dipole field, the line corresponding to λ=55° intersects the plane of the geomagnetic equator at 2.04 earth radii, or 13,010 km above the earth's surface. (Note that fig. 4 shows the model only out to the height 3,500 km.) The total length along the line of force from the equator to the earth's surface is 3.17 earth radii, or 20,214 km. At the equator the magnetic field strength along the line of force is a minimum. Its value is 0.011 gauss compared to 0.315 gauss at the surface. The corresponding gyrofrequency is 31.4 kc, and with the exosphere model (7) N_e=1,560 and f_p=354 kc at this height.

Using as the origin of time a starting position in the equatorial plane, the calculations were made for points regularly spaced along the trajectory to obtain: (1) the frequencies emitted; (2) the time of the emission; (3) the time for propagation of the wave from this point to the earth's surface along the remaining arc, using the full formula for group velocity in the whistler mode; and, (4) the sum of these two times to obtain the arrival time at the earth.

At 10,000 km/sec the total time of flight for the cloud is nearly 2 sec while the range of arrival times of the emitted waves is about ½ sec. Of course, one observes only the range of arrival times.

The "knee" in the "hook" in the present example corresponds to the transition at about 2,500-km height between the electron density characteristic of the exosphere and the more rapidly increasing density in the ionosphere, which is responsible for the high branch of the "hook."

7.4. Velocity of the Cloud

Generally speaking, the shape of the computed spectrum will remain approximately the same for different velocities. But the frequency and time scales will vary rapidly. A small velocity furnishes a range of arrival times that are too long and emitted frequencies that are too low. The durations and frequencies are very sensitive to the velocity. Velocities of about 1,000 km/sec, as suggested by the delay between related solar and terrestrial events, give frequencies roughly ten times too small and durations 10 times too long. Satisfactory fits are obtained only for velocities of the order of 10,000 km/sec. Such a high velocity seems to indicate that some process in the interaction between the earth's exosphere and the solar corpuscular radiations gives high local velocities to the small clouds responsible for the vlf emissions. These clouds are tentatively associated with the flashes of luminosity in active auroras.

It is significant that this velocity is of the same order as that required for protons to penetrate to the 100-km level in the aurora.

TABLE 2. *Summary of the conditions assumed for the sample theoretical calculation*

Distribution of electron density		Other conditions
Ionosphere	Exosphere	
$N_{max}=10^6/cm^3$	$h_{ex}=1{,}000$ km	$\lambda=55°$
$h_{max}=350$ km	$\mathscr{H}_{ex}=200$ km	$v_e=10^4$ km/s
$\mathscr{H}_A=200$ km	$K=4{,}000$ kc	
$\mathscr{H}_B=100$ km		

The work reported in this paper was supported partially by the Air Force Office of Research under Contract AF18 (603)–126. The helpful suggestions of T. N. Gautier and L. R. O. Storey are gratefully acknowledged.

8. References

[1] L. R. O. Storey, Ph. D. dissertation, Cambridge University, England, and Phil. Trans. Roy Soc. [A]**246,** 113 (1953).
[2] R. A. Helliwell, J. H. Crary, J. H. Pope, and R. L. Smith, The nose whistler, a new high latitude phenomenon, J. Geophys. Research **61,** 139 (1956).
[3] G. R. Ellis, J. Atmospheric and Terrest. Phys. **8,** 338 (1956).
[4] K. Maeda and I. Kimura, A theoretical investigation on the propagation path of the whistling atmospherics, Rep. Ionosphere Research Japan **10,** 105 (1956).
[5] J. H. Pope, Diurnal variation in the occurrence of Dawn Chorus, Nature **180,** 433 (1957).
[6] G. McK. Allcock, A study of the audio-frequency radio phenomenon known as Dawn Chorus, Australian J. Phys. **10,** 286 (1957).
[7] H. E. Dinger, Whistling atmospherics, Nav. Research Lab. Rep. 4825 (September 14, 1956).
[8] J. M. Watts, An observation of audio-frequency electromagnetic noise during a period of solar disturbance, J. Geophys. Research **62,** 199 (1957).
[9] G. R. Ellis, Low frequency radio emission from aurorae, J. Atmospheric and Terrest. Phys. **10,** 302 (1957).
[10] J. R. Pierce, Traveling-wave Tubes (D. Van Nostrand Co., New York, N.Y., 1950).
[11] A. B. Meinel, Evidence for the entry into the upper atmosphere of high-speed protrons during auroral activity, Science, **112,** 590 (1950).
[12] L. Vegard, Terrestrial magnetism and electricity, J. A. Fleming, ed., p. 609 (McGraw-Hill, New York, N.Y., 1937).
[13] Stanford University, Low-frequency propagation studies, pt. 1, Whistlers and related phenomena, Final report, Contract AF 19 (604)–795 (June 15, 1953 to September 30, 1956).
[14] D. F. Martyn, The theory of magnetic storms and auroras, Nature, **167,** 92 (1951).

BOULDER, COLO. (Paper 63D1–4).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Climatology of Ground-Based Radio Ducts*

Bradford R. Bean

(January 15, 1959)

An atmospheric duct is defined as occurring when geometrical optics indicate that a radio ray leaving the transmitter and passing upwards through the atmosphere is sufficiently refracted that it is traveling parallel to the earth's surface. Maximum observed incidence of ducts was determined as 13 percent in the tropics, 10 percent in the arctic and 5 percent in the temperate zone by analysis of 3 to 5 years of radiosonde data for a tropical, temperate, and arctic location. Annual maximums are observed in the winter for the arctic and summer for the tropics. The arctic ducts arise from ground-based temperature inversions with the ground temperature less than −25° C while the tropical ducts are observed to occur with slight temperature and humidity lapse when the surface temperature is 30° C and greater.

1. Introduction

The author has recently had cause to investigate the limitations placed upon ray tracing of vhf-uhf radio waves by the occurrence of atmospheric ducts [1].[1] Ducting is defined as occurring when a radio ray originating at the earth's surface is sufficiently refracted during its upward passage through the atmosphere so that it either is bent back towards the earth's surface or travels in a path parallel to the earth's surface. Although the proper treatment of ducting involves consideration of the wave equation solution [2] rather than a simple ray treatment, the present study will be based upon a geometrical optics definition of the limiting case in which ray tracing techniques may be used. This simple criterion is then applied to several years of radiosonde observations from stations typical of arctic, temperate and tropical climates to derive estimates of the variation of the occurrence of radio ducts with climatic conditions.

2. Background

The property of the atmosphere basic to radio-ray tracing is the gradient of the radio refractive index of the atmosphere, n. For standard conditions near the surface of the earth n is a number of the order 1.0003 and its gradient is about 40×10^{-6} per kilometer. It is instructive to consider the order of magnitude of refractive index gradient needed for trapping for several commonly observed refractive index profiles. Snell's law may be written, for cylindrical coordinates,

$$n_t r_t \cos\theta_t = n_d r_d \cos\theta_d, \tag{1}$$

where θ is the elevation angle made by the ray at the point under consideration. The subscripts t and d refer to the values of the variables at the transmitter height and the top of the trapping layer respectively. Thus trapping occurs when the ray is traveling parallel to the earth, i.e., $\cos\theta_d=1$ and

$$\frac{n_t r_t}{n_d r_d} \geq 1. \tag{2}$$

The angle of penetration at the transmitter, θ_p, found by setting

$$\frac{n_t r_t}{n_d r_d}\cos\theta_p = 1, \tag{3}$$

divides the ray family into two groups since all rays of $\theta_0 \leq \theta_p$ are trapped within the duct and those rays of $\theta_0 > \theta_p$ are not. The n gradient for a given value of θ_p is then given by

$$\frac{\Delta n}{\Delta r} = -\frac{n_t - n_d}{r_d - r_t}, \tag{4}$$

where, for the ducting case, n_d must satisfy (3), i.e.,

$$n_d = \frac{n_t r_t}{r_d}\cos\theta_p. \tag{5}$$

By designating $r_d = r_t + \Delta h$ then (4) may be written

$$\frac{\Delta n}{\Delta r} = -\frac{n_t}{\Delta h}\left[1 - \frac{r_t}{r_t + \Delta h}\cos\theta_p\right]. \tag{6}$$

By rewriting (6),

$$\frac{\Delta n}{\Delta r} = -\frac{n_t}{\Delta h}\left[1 - \frac{1}{1+\frac{\Delta h}{r_t}}\cos\theta_p\right], \tag{7}$$

and expanding $\left(1+\frac{\Delta h}{r_t}\right)^{-1}$ and $\cos\theta_p$ one obtains the expression

$$\frac{\Delta n}{\Delta r} \cong -n_t\left[\frac{1}{r_t} + \frac{\theta_p^2}{2\Delta h}\right] \tag{8}$$

by neglecting terms of the order $\frac{\theta_p^4}{4!}$ and $\left(\frac{1}{r_t}\right)^2$.

* This work was partially sponsored by task 31 of the U. S. Navy Weather Research Facility, Norfolk, Va.

[1] Figures in brackets indicate the literature references at the end of this paper.

For the case of $\theta_p=0$ and the transmitter at sea level, (8) reduces to

$$-\frac{\Delta n}{\Delta r}=\frac{n}{a}\cong\frac{1}{a}\cong 157\,N \text{ units/km}. \tag{9}$$

using $a=6{,}373$ km and $N=(n-1)\ 10^6$. Note that the n gradient is referred to here and hereafter as parts per million, or in popular parlance, N units.

It is seen from (8) that the n gradient necessary to trap a radio ray at a given value of θ_p is practically independent of transmitting antenna height above the earth. For example, a $\theta_p=0$ ray will be trapped by an n gradient of -157.0 N units/km at sea level when $n_t=1.0003$ while the necessary n gradient at 3 km above sea level will be -156.9 N units/km for an $n_t=1.0002$, thus indicating, for all practical applications, that the necessary n gradient for trapping is independent of altitude. Further, by considering the temperature and humidity gradients encountered in the troposphere one is led to the conclusion that ducting gradients would not be expected to occur at heights greater than 3 km above the earth's surface. In fact, Cowan's [3] investigation indicates that trapping gradients are nearly always confined to the first kilometer above the surface.

A consideration of (8) indicates that the magnitude of the negative gradient necessary for ducting is $1/a$ for $\theta_o=0$ but is increased by the amount $n_t\theta_p^2/2\Delta h$ for other values of θ_p. The gradients necessary for atmospheric ducts as a function of θ_o are given for several different but typical n profiles in figure 1. An analysis of radiosonde data indicates that gradients in excess of 500 N units/km are seldom indicated within atmospheric layers. Note how rapidly the necessary gradients increase to the approximate upper limit of gradients derived from radiosonde observations; a ground-based layer 100-m thick attains this gradient at 8.3 milliradians while the maximum observed gradient is intercepted by the 30-m layer curve at 4.5 milliradians of θ_o. A third example was calculated for an elevated layer 0.5 km above the ground and 100-m thick by assuming normal refraction between the ground and the base of the layer and solving for the necessary ducting gradient within the layer. The large values of the n gradient necessary for this case indicate that elevated ducts would rarely be observed. Although the preceding examples were calculated for a ground transmitter, the combinations of θ_p, $\Delta n/\Delta h$, and Δh are very nearly the same as would be obtained for any other transmitter height within the first 3 km above the surface.

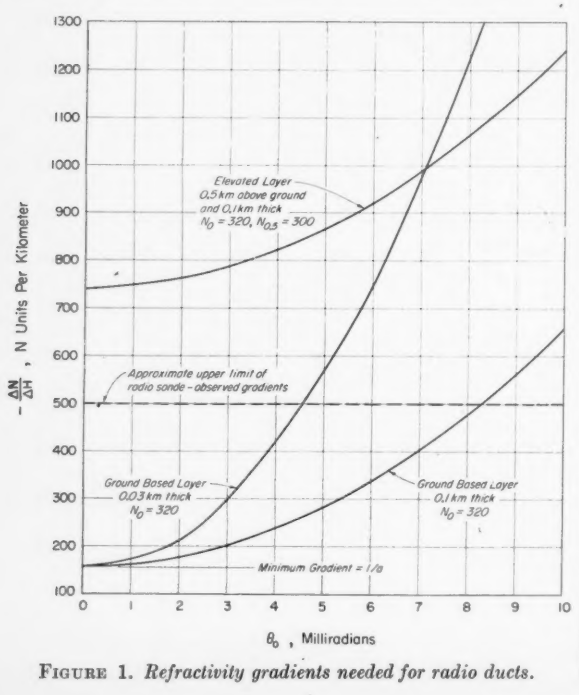


FIGURE 1. *Refractivity gradients needed for radio ducts.*

3. Description of Observed Ground-Based Atmospheric Ducts

Radiosonde data were examined for the occurrence of ducts. Three consecutive years of data were analyzed for the months of February, May, August, and November at each of three Weather Bureau stations. The three stations were chosen to represent a range of climates: Fairbanks, Alaska for an arctic climate, Washington, D. C., for a temperate climate and Swan Island, West Indies, as an example of a tropical climate. The procedure used to determine the occurrence of a radio duct was to:

(a) Determine the value of N from the expression [4]

$$N=(n-1)10^6=\frac{77.6}{T}\left(P+\frac{4810\,e_s RH}{T}\right), \tag{10}$$

where P is the station pressure in millibars, RH is the percent of the saturation vapor pressure, e_s, in millibars at the absolute temperature, T, in degrees Kelvin;

(b) note all instances when the N gradient equaled or exceeded the minimum ducting gradient indicated by (9), i.e.,

$$-\frac{\Delta n}{\Delta r}=\frac{n}{a}\sim\frac{1}{a}=157\,N \text{ units/km}; \tag{11}$$

(c) if, for the instances selected by (b),

$$\frac{n_t r_t}{n_d r_d}\geq 1, \tag{2}$$

where $r_t=a$, then the duct was said to trap rays from an antenna resting on the surface of the earth. Further, under this condition the particular duct would trap all rays from $\theta_o=0$ up to the angle of penetration,

$$\cos\theta_p=\frac{n_d r_d}{n_t a}. \tag{12}$$

The ducts selected by procedure (a) through (c) are defined as ground-based ducts. Statistics of ground-based ducts are given below.

The percentage occurrence of ducts is shown on figure 2. The maximum occurrences of 13.8 percent for August at Swan Island and 9.2 percent for Fairbanks in February are significantly greater than the values observed at other times of the year. The Washington data display a summertime maximum of 4.6 percent indicating the temperate zone maximum incidence is about one-half the wintertime maximum incidence in the arctic, and about one-third of the summertime tropical maximum.

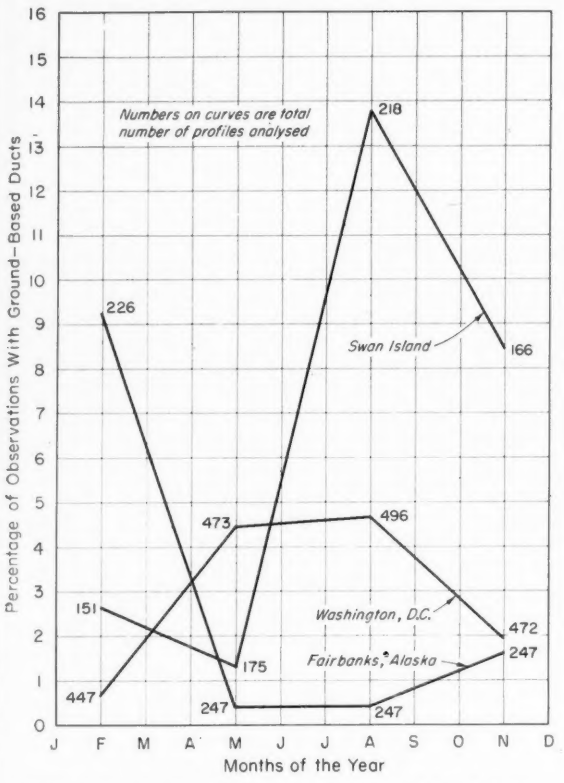


FIGURE 2. *Frequency of occurrence of ground-based ducts.*

The range of values of θ_p observed are shown in figure 3. The mean value calculated for each month as well as the maximum and mean values of θ_p observed for the limiting cases are given for each month and location. The mean value of the angle of penetration under these conditions is between 2 and 3 milliradians and appears to be independent of climate. The maximum value of θ_p observed during ducting is 5.8 milliradians.

The refractivity gradients observed during ducting are given on figure 4. The maximum gradient of 420 N units/km was observed during February at Fairbanks, Alaska. The mean values of N gradient appear to follow a slight climatic trend from a high

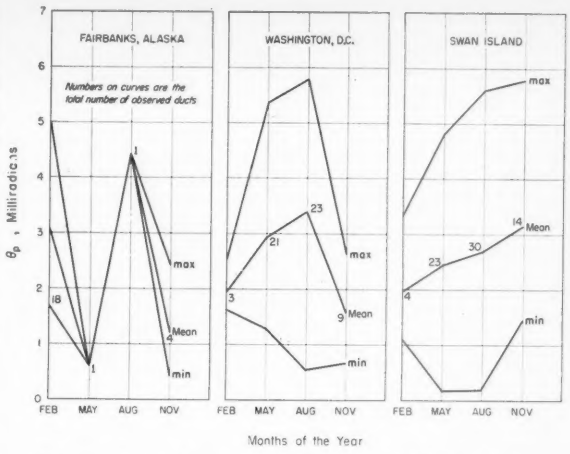


FIGURE 3. *Angle of penetration of ground-based ducts.*

value of 230 N units/km at Fairbanks to a value of 190 N units/km at Swan Island.

Another property of radio ducts is their thickness which is given in figure 5. Again there is observed a slight climatic trend as the median thickness increases from 66 m at Fairbanks to 106 m at Swan Island. These values of thickness correspond to the gradients given in figure 4. The steepness of the distribution curve for Washington appears to be due to a mixture of N profile types that give rise to ground-based ducts. The thickest ducts at Washington are actually elevated layers accompanied by a relatively small N gradient between the ground and the base of the layer. This is in contrast to the Fairbanks and Swan Island profiles which tend to be composed of a single ducting gradient from the surface to the top of the duct.

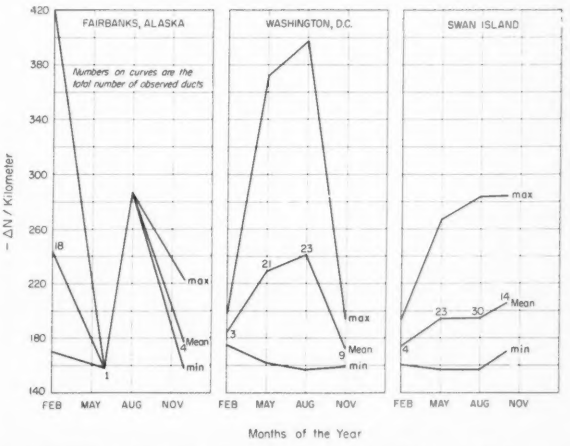


FIGURE 4. *Refractivity gradients of ground-based ducts.*

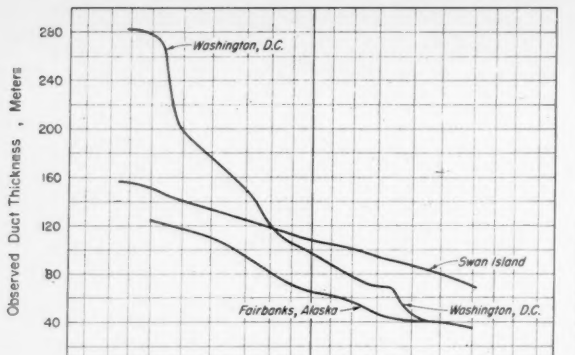


FIGURE 5. *Observed ground-based duct thickness.*

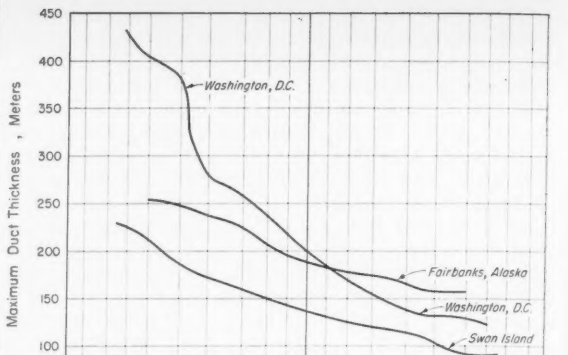


FIGURE 6. *Maximum ground-based duct thickness.*

One may obtain yet another thickness by linearly extrapolating to obtain the height at which the gradient is equal to $-1/a$, that is, the height corresponding to the gradient just sufficient to trap the ray at $\theta_o=0$. These values, shown in figure 6, display an increase in the median thickness of about 25 percent for Swan Island, 100 percent for Washington, and 200 percent for Fairbanks, which results in a reversal of the climatic trend of the observed thickness between Fairbanks and Swan Island. This increase in height emphasizes the information of the preceding figures, namely, Fairbanks is characterized by shallow layers with relatively intense gradients.

These maximum duct widths may be used to estimate the minimum frequencies that are trapped by reference to a ducting theory that assumes a linear decay of refractive index within the duct such as that given by Kerr [5] where the maximum wavelength, λ, trapped by a given thickness, d, is given by

$$\lambda_{max}=c\gamma^{1/2}d^{3/2}, \tag{13}$$

where c is a constant and γ is a function of the n gradient excess over the minimum value of $\Delta N/\Delta H=1/a$. If λ_{max} is to be expressed in centimeters, d in meters and

$$\gamma=\left(\frac{N_t-N_d}{d}-15.7\right)10^{-8}, \tag{14}$$

then

$$c=2.514\times10^2.$$

By the use of (13) the minimum frequencies trapped during ducting conditions were estimated for the maximum duct thicknesses of figure 6. These values, given in table 1, were determined for the month with the maximum occurrence of ducts, thus allowing an estimate of the radio frequencies likely to be effected by ducting conditions. Note, for example, that the data of table 1 indicate that 500-Mc rays will be trapped by 50 percent of the ducts regardless of location.

The reader is cautioned that an atmospheric duct does not have the sharp boundaries of a metallic waveguide and thus the minimum frequencies given by table 1 do not represent cutoff frequencies but, as Kerr is so careful to emphasize, merely yields a suggested lower limit of the frequencies strongly affected by the duct under the assumptions of this rudimentary theory.

TABLE 1. *Estimated minimum frequency trapped at* $\theta_o=0$

Station	Minimum frequency in megacycles trapped by the indicated percentage of ducts						
	95%	*90%*	*75%*	*50%*	*25%*	*10%*	*5%*
Fairbanks, Alaska (February)	1,500	1,300	1,200	890	690	490	435
Washington, D.C. (August)	4,300	3,000	1,100	600	270	180	150
Swan Island, West Indies (August)	2,500	1,300	725	500	365	325	225

4. Temperature and Humidity Distributions Associated with Ground-Based Ducts

The temperature and humidity structure within ground-based ducts appears to be somewhat different from that normally encountered in the atmosphere. This departure from the normal structure is an aid in distinguishing the different atmospheric mechanisms that give rise to ducts as well as helping the meteorologist forecast ducting from his experience with the normal meteorological variables.

The temperature and humidity structure of the atmosphere during ducting conditions may be evaluated by noting that N is composed of a term proportional to the air density, D, plus a term involving the partial pressure of water vapor, W.

These components are given by

$$D=\frac{77.6}{T}P, \tag{15}$$

and

$$W=\frac{3.73\times10^5 e_s RH}{T^2}. \tag{16}$$

The gradient, ΔN, with respect to height may then be expressed:

$$\Delta N=\Delta D+\Delta W. \tag{17}$$

Average values of ΔN, ΔD, and ΔW are given for two increments between the earth's surface and 1 km above sea level for Fairbanks, Alaska, Washington, D.C., and Swan Island, W.I., in table 2. These data were determined from the Weather Bureau publication on long term mean upper air data [6].

TABLE 2. *Gradient of* N, D, *and* W (*N units/km*)

Station	Height increment	February			August		
		$-\Delta N$	$-\Delta D$	$-\Delta W$	$-\Delta N$	$-\Delta D$	$-\Delta W$
Fairbanks, Alaska	surface to 0.5 km	37	41	−4	31	27	4
	0.5 km to 1.0 km	35	35	0	36	24	12
Washington, D.C.	surface to 0.5 km	41	34	7	60	28	32
	0.5 km to 1.0 km	30	26	4	46	24	22
Swan Island, West Indies	surface to 0.5 km	39	24	15	47	26	21
	0.5 km to 1.0 km	58	24	34	66	24	42

Several general observations may be made of the data of table 2. The gradient of the dry term is relatively less variable than that of the wet term when considered as a function of season or height; the increase of ΔN from winter to summer at a particular location or from arctic to tropical climate at a given time is most strongly reflected in ΔW rather than in ΔD. The marked increase of gradient with height for Swan Island reflects the drop of refractivity across the interface of the trade wind inversion where dry subsiding air overlies the moist oceanic surface layer. Note, however, that in all cases the average N gradient is significantly less than the -157 N units/km needed for ducting.

Examples of the temperature and humidity structures within ducts typical of each of our three climates are shown in figure 7. The Fairbanks duct is accompanied by a surface temperature of −25° C with a strong temperature inversion and a slight humidity lapse indicating temperature inversions associated with wintertime cooling of the air next to the ground. The Washington example appears to be typical of the temperate-zone temperature inversion. The Swan Island profile, however, shows a moderate lapse of both temperature and humidity. This apparent contradiction is explained by the strong lapse in vapor pressure associated with the moderate lapse in temperature when the initial temperature is near 30° C. The strong vapor pressure lapse presumably arises from evaporation off the sea surface. This effect was further examined by studying the percentage of the total N gradient of each duct that was contributed by the gradient of the dry and wet terms. The median contribution of the dry term gradient, summarized in table 3, displays strong seasonal and geographic differences. The dry term contribution decreases from summer to winter and from arctic to tropical climates. The Swan Island ducting gradients are at least 90 percent due to humidity lapse, while the Fairbanks wintertime maximum is due to the strong temperature inversion associated with very low surface temperatures. In fact, under these conditions at Fairbanks the vapor pressure actually increases with height with the result that the dry term contributes more than 100 percent of the ducting gradient.

TABLE 3. *Median contribution of* $\Delta D/\Delta H$ *to* $\Delta N/\Delta H$ *for ducting conditions*

	Fairbanks	Washington, D.C.	Swan Island
	%	%	%
February	103.0	73.0	9.5
May	40.5	33.5	2.0
August	37.0	26.5	4.5
November	62.0	55.0	6.0

The data for Washington, D. C., however, appears to indicate that temperate zone ducting arises from a mixture of the arctic and tropic mechanisms dependent upon the season. It appears that the wintertime ducts in the temperate zone are of the dry-term arctic type while the summertime ducts are of the tropical humidity-lapse variety.

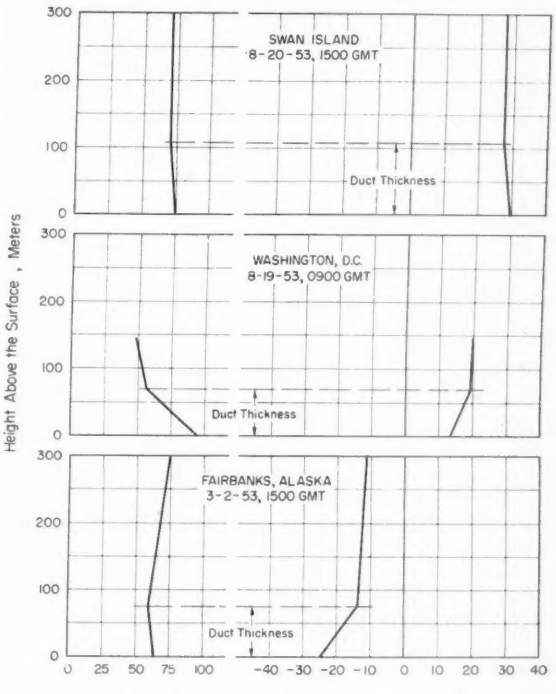


FIGURE 7. *Temperature and humidity profiles for typical ground-based ducts.*

5. Conclusions

The results given above were derived from a consideration of radiosonde data. Although the radiosonde is not an extremely sensitive instrument, it is readily available and forms the only source of climatic information involving the temperature and humidity structure of the atmosphere. It is believed that the radiosonde data will at least yield information as to the climatic trend of radio ducts as well as their probable temperature and humidity distributions. Further, it is evident that the choice of stations may affect the actual percentage of ducts observed. For example, it is certain that a greater percentage of ducts would be observed over water in the subtropics than over Swan Island. Indeed, the kite soundings of the Meteor Fxpedition [7] indicate about a 50-percent incidence of *ducting gradients* (but not necessarily ground-based ducts) between 6° and 20° N latitude and between 9° and 30° S latitude.

With these reservations in mind, the present study has shown:

(a) Limiting layers occur no more than 15 percent of the time.

(b) The annual cycle of the incidence of limiting layers is reversed for the arctic and tropical stations studied. The arctic station has a wintertime maximum and the tropic station a summertime maximum. The temperate station has a summertime maximum incidence of less than 5 percent.

(c) The maximum angle of penetration observed is 5.8 milliradians with a mean value of about 3 milliradians.

(d) The maximum observed lapse rate of N is -420 N units/km.

(e) The maximum thickness of observed limiting layers is such as to trap radio waves with frequencies greater than about 500 Mc at all locations for at least 50 percent of the observed ducts.

(f) The limiting layers in the arctic appear to be associated with temperature inversions at ground temperatures of $-25°$ C or less; temperate zone, with the common radiation inversion and accompanying humidity lapse; tropics, with a moderate temperature and humidity lapse for temperatures of about 30° C.

One may wonder how much the above conclusions are affected by the choice of the antenna height $h_t=0$. This question may be partially answered by a consideration of the climatic occurrence of ducting gradients with respect to height. Cowan's analysis of ducting gradients has shown that such gradients may be found in as many as 40 percent of the observations during some months at some locations. This same study indicates that the majority of ducting gradients are observed within the first kilometer above the earth and that the incidence of ducting gradients decreases rapidly with increasing height becoming negligible for heights in excess of 3 km. The inference of these data is that the ground-based duct statistics are representative of maximum conditions and that the effects of ducts will be negligible above 3 km.

If, on the other hand, a climate is such that elevated ducts consistently form at some altitude, then the choice of antenna height would definitely bias the ducting statistics. For example, during the summer at San Diego and Oakland, Calif., an elevated duct is observed for 37 to 54 percent of the observations [8] with an average base 0.55 km above the surface, an average thickness of 0.275 km and an average gradient of 195 N units/km. This elevated duct would not trap any rays if $h_t=0$. If, however, the antennas were located at the base of the duct ($h_t=$ 0.55 km) then all rays of $\theta_o \leq 12$ milliradians would be trapped. The angle of penetration would become less with increasing distance of the antenna below the base of the duct becoming zero when h_t is 375 m below the duct base.

In any case encountered in practice, however, one should either determine the duct incidence for any particular combination of antenna height and climate or consider the above statistics to be indicative of the relative climatic distribution of radio ducts.

The author expresses his gratitude to G. D. Thayer for the preparation of an electronic digital computer program, to Mrs. B. J. Weddle and Miss M. F. Auer for aiding in the calculations of this paper, and to Miss G. Hoffmire for the preparation of the manuscript.

6. References

[1] B. R. Bean and B. A. Cahoon, The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles, Proc. IRE **45,** 1545 (1957).

[2] H. G. Booker and W. Walkinshaw, The mode theory of tropospheric refraction and its relation to wave guides and diffraction, Meteorological factors in radio wave propagation, p. 80, Phys. Soc. and Roy. Meteorol. Soc. conference report (London, Eng., April 8, 1946).

[3] L. W. Cowan, A. radio climatology survey of the U.S., Proc. Conf. on Radio Meteorol. (Univ. of Texas, Austin, Texas, Nov. 9–12, 1953).

[4] E. K. Smith and S. Weintraub, The Constants in the equation for atmospheric refractive index at radio frequencies, Proc. IRE, **41,** 1035 (1953).

[5] D. E. Kerr, Propagation of short radio waves, p. 9 (McGraw-Hill Book Co., Inc., New York, N.Y., 1951).

[6] B. Ratner, Upper air average values of temperature, pressure, and relative humidity over the United States and Alaska (U.S. Weather Bureau, 1945).

[7] D. E. Kerr, Propagation of short radio waves, p. 262 (McGraw-Hill Book Co., Inc., New York, N.Y., 1951).

[8] D. E. Kerr, Propagation of short radio waves, p. 261 (McGraw-Hill Book Co., Inc., New York, N.Y., 1951).

Boulder, Colo. (Paper 63D1–5).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Power Requirements and Choice of an Optimum Frequency for a Worldwide Standard-Frequency Broadcasting Station

A. D. Watt and R. W. Plush

(January 15, 1959)

Calculations are presented for the expected transmission characteristics and atmospheric noise levels in the 8- to 100-kc band. When these are combined with carrier-to-noise requirements for a given precision of frequency comparison, it is indicated that a minimum radiated power in the order of 10 to 100 kw for frequencies in the vicinity of 20 kc will be required to provide worldwide coverage. Minimum observation times of 15 to 30 min appear to be required for these transmitter powers in order to obtain a precision of frequency comparison of 1 part in 10^9 for typical transmission paths. Carrier-to-noise requirements and the factors determining this ratio are considered for typical receiving systems.

1. Introduction

Observations by various investigators of the frequency stability of lf and vlf radio signals at great distances has led to the proposal by J. A. Pierce, et al. [1][1] for a single standard frequency transmitter with worldwide coverage. Descriptions of various radio systems employed for comparing frequency are given by Pierce [2], along with the stabilities obtainable. Other observations on frequency stability are reported by Allan, Crombie, and Penton [3], and some of the possible advantages of a vlf or lf frequency standard over the present high frequency transmissions from WWV are discussed by George [4].

It is the purpose of this investigation to make preliminary determinations of the optimum frequency and the power requirements for reliable coverage. The first objective will be to determine the radiated power required to provide this coverage as a function of frequency. Since the atmospheric noise level is well in excess of antenna or receiver thermal noise in this frequency range for reasonable sized antennas, we can write

$$P_r = -E_1 + E_{nm} + C/N_{1\,\mathrm{kc}} + T_x. \qquad (1)$$

All factors are expressed in db: P_r is the radiated power relative to 1 kw, E_1 is the field produced at the receiving location for 1 kw radiated relative to 1 μv/m, E_{nm} is the median rms noise field in a 1 kc band relative to 1 μv/m, $C/N_{1\,\mathrm{kc}}$ is the required rms carrier to rms noise in a 1 kc effective bandwidth for the type of service involved, and T_x is a factor which assures this type of service for a given percentage of all hours in spite of the time variability of the noise as well as the variation in carrier field strength due to propagation effects.

[1] Figures in brackets indicate the literature references at the end of this paper.

2. Carrier Field Intensity Calculations

Before calculating E_1 the field intensity anticipated for 1 kw radiated, we must first determine the distance to the areas most difficult to serve. With the aid of a globe and noise maps by Crichlow, et al. [5], it readily becomes apparent that for a transmitter located near the Boulder Laboratories in Colorado the most difficult service areas in terms of carrier-to-noise requirements will be in the vicinity of Java or Madagascar. The distance to these areas is approximately 17,000 km. The transmission paths are over both land and sea water, and this should be allowed for in the final calculations.

The large amount of recent theoretical and experimental work in the field of vlf propagation [6] has placed us in a position where relatively accurate predictions of field intensity can be made in this frequency range. Studies by Wait [7] have presented expressions for the vertical electric field which can be well approximated for ranges in excess of 2,000 km as

$$E \approx K + P_r - 10\log_{10} f(\mathrm{kc}) - 10\log_{10}[a\sin(d/a)] - \alpha d/1{,}000 \qquad (2)$$

$d > 2{,}000$ km,

where E is the vertical electric field in db relative to 1 μv/m at a distance d from the source, d is this distance in kilometers, K is a constant (see appendix A) which equals 97.5 for day paths assuming an ionospheric height of 70 km and equals 94.8 for night paths where $h=90$ km, P_r is the radiated power in db relative to 1 kw, f(kc) is the frequency in kc, a is the earth's radius ($\sim$6,400 km), and α is the attenuation rate in db per 1,000 km. All except the last term represent the unabsorbed field expected from the dominant mode being propagated between two concentric spherical shells. The term α, which

accounts for the fact that the shells are not perfectly conducting, will naturally vary with ionospheric conditions, ground conductivity, frequency, and perhaps to a small extent on direction of propagation with respect to the earth's magnetic field. Experimental determinations of the factor α can be made either by employing equation (2), or by observing the rate of decrease in E with distance relative to the $(a \sin d/a)$ factor. Determinations of both types have been made employing observations by Round et al. [8], Pierce [9] and Heritage [10], along with theoretical calculations by Wait [11] for both single and double layer ionospheric models, and the results presented in figure 1. It should be emphasized that

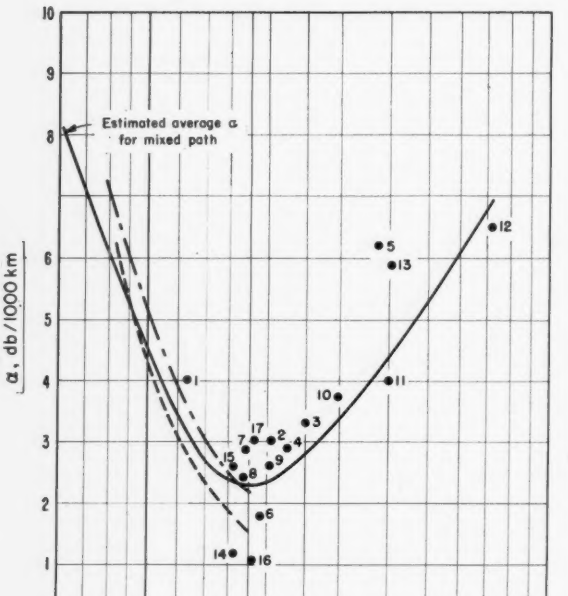


FIGURE 1. *Attenuation coefficient for daytime conditions*

No.	Author Ref.	Transmitter	Cosχ* Receiver	Path
1	Round, et al. [8]	LY	Dorset	L & S
2	do	POZ	do	L & S
3	do	LCM	do	L & S
4	do	KET	do	L & S
5	do	GB	do	L & S
6	do	YN	Boonah	L & S
7	do	LY	do	L & S
8	do	UFU	do	L & S
9	do	OUI	do	L & S
10	Pierce [9]	San Diego 0.96	Hawaii	S
11	do	do .96	do	S
12	do	do .96	do	S
13	do	do .93	Wash., D.C.	L
14	Heritage [10]	NPM	Aircraft	S
15	do	NPM	do	S
16	do	NLK	do	S
17	do	NLK	do	S
	Wait [11]	Theory	ground $\{\sigma=\infty$	
	do	do	$\sigma=2.2$ mmho/m	

L=Land
S=Sea
*Max. sun's zenith path midpoint

these attenuation rates are only approximate median values for low latitude paths with relatively few discontinuities and a fairly low percentage of land in the total path. Unfortunately, present knowledge in this field is not sufficient to permit extremely accurate predictions of α for all paths. It does appear, however, that α is considerably greater for high latitude paths when appreciable lengths of low conductivity earth surface such as permafrost or glaciers are involved. In addition, the total loss over a path appears to increase when appreciable discontinuities in surface conductivity and/or height of the ionosphere are present over the path. This last effect which occurs when a sunrise or sunset boundary is crossed by the path may cause an effective decrease in carrier field by approximately 6 to 10 db and 3 to 6 db respectively over some but not necessarily all paths.

Since α varies appreciably with ionospheric conditions, our estimated α in figure 1 is only expected to be near some median value. Experimental substantiation of the general shape of this curve in the low frequency region is given by atmospheric noise measurements [12] and by Taylor [13] in his spherics observations.[1] It should be pointed out that the nighttime attenuation values are expected to be more variable than the daytime values, and in addition α will not increase as rapidly with frequency above 18 kc at night as is shown for the day conditions.

The values of α from figure 1 are now employed in eq (2) for a 17,000-km path and an assumed radiated power of 1 kw. The results are shown in figure 2 where it is evident that at this distance the propagation path appears to have a rather narrow pass band centered around 18 kc.

So far we have not considered the affects near the antipode where the signal is arriving from essentially all directions. In theory, at this point the electric field focuses and an appreciable build up of some 20 db at 20 kc is possible. Observations in this area by Round, et al. [8] have shown that at times the measured field can be 20 db above that anticipated for a single path at this distance. Their observations also indicate, as would be expected, that beats are produced and that the received field is not very stable due to interference between the various paths. Some of this interference effect can possibly be reduced by loop and vertical antenna combinations or perhaps some type of angle diversity with loops. Fortunately, for a transmitting site near Boulder, the antipode is in the Indian Ocean where little use of the standard frequency broadcast service is presently anticipated.

The possibility of whistler mode propagation producing interference as shown by Helliwell [23] should also be considered; however, again the conjugate point from Boulder is not in an area of great importance.

[1] Additional material by Wait [25] on attenuation rates including an extensive bibliography has become available after this study was completed. In general, the attenuation rates indicated by Wait are in good agreement with figure 1, although it is evident that further careful experimentation is desired to determine with greater precision the shape of this curve and the frequency of minimum attenuation.

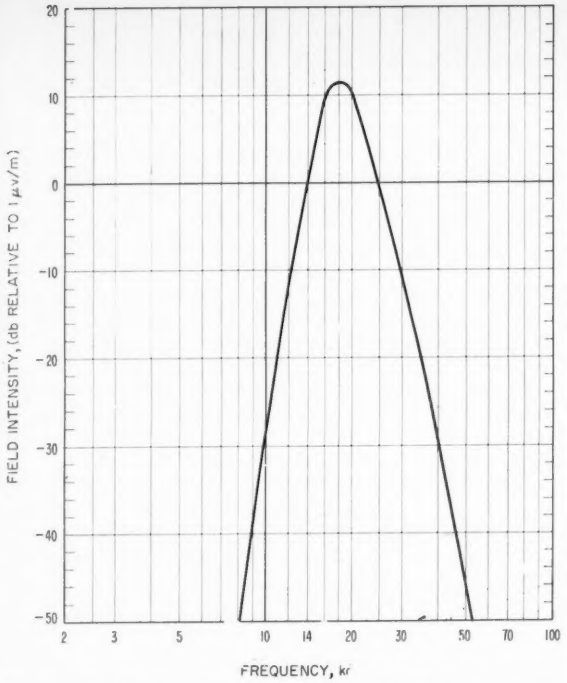


FIGURE 2. *Field intensity expected at 17,000 km for 1 kw radiated power.*

Daytime, mixed land and sea path: land ≈ 20 percent, sea ≈ 80 percent (average ionospheric conditions assumed). Calculations based on eq 2 and α of figure 1.

3. Atmospheric Noise Levels

The second term in eq (1), E_{nm} describes the median rms noise field in a 1-kc bandwidth, and values anticipated in the regions considered can be found in CCIR Report No. 65 [5]. The maximum level anticipated during the summer months corresponds to a noise grade of 90 db at 1 mc, and the anticipated median values in the frequency range considered are shown in figure 3. Since the high noise level results from local thunderstorm activity, and the predictions are extrapolated to 10 kc from higher frequency measurements, the shape of this curve has been slightly modified to agree with the average spectrum of the radiation component from lightning strokes [12]. In addition, several points have been included from an earlier paper [14] along with a typical winter curve [12].

4. Carrier-to-Noise Requirements

The ratio of rms carrier to effective rms noise in a 1-kc bandwidth, $C/N_{1\text{ kc}}$, is useful in defining the amount of carrier power required for a satisfactory determination of frequency. Pierce [2] has shown that several different types of equipment are presently employed for frequency comparison, and the actual $C/N_{1\text{ kc}}$ required will undoubtedly vary with the characteristics of the receiving equipment and

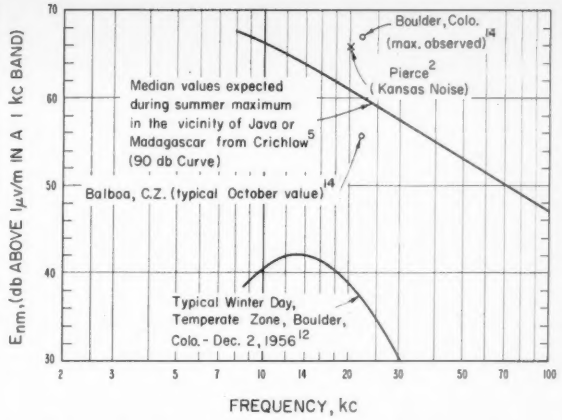

Boulder, Colo. (max. observed) [14]
Pierce [2] (Kansas Noise)
Median values expected during summer maximum in the vicinity of Java or Madagascar from Crichlow [5] (90 db Curve)
Balboa, C.Z. (typical October value) [14]
Typical Winter Day, Temperate Zone, Boulder, Colo. - Dec. 2, 1956 [12]
E_{nm}, (db ABOVE 1μv/m IN A 1 kc BAND)
FREQUENCY, kc

FIGURE 3. *Typical atmospheric noise field strengths.*

Rms values in a 1-kc effective bandwidth.

the amount of effort spent in attempting to separate the desired energy from the undesired noise energy.

It is well known that the carrier-to-noise ratio required to make a frequency comparison is dependent on the precision required and the period of time allowed to make the frequency comparison. The precision of frequency comparison is defined as $\epsilon = \sigma_{fd}/f$ where f is the standard frequency, and σ_{fd} is the standard deviation of the frequency difference between the resultant received frequency and the standard frequency.[2] Since there is a definite minimum time required for a given precision of frequency comparison due to propagation path phase instabilities, we must consider this effect first. This effect is analyzed in appendix B where it is shown that the ϵ obtainable is independent of observing time for periods short compared to the carrier fade period and then becomes inversely proportional to this period for some systems. It also appears that for periods long compared to the carrier fading that some systems provide a decrease in ϵ which is proportional to the 3/2 power of time. Calculations in the 20-kc region indicate that frequency comparisons can be made with a precision of one part in 10^9 in observing periods of 15 to 30 min.

The factors which determine required $C/N_{1\text{ kc}}$ are considered in appendix C. The results are shown in figure 4 which gives anticipated values of $C/N_{1\text{ kc}}$ for several observing times. It is interesting to note here that if the carrier is required to carry information at an appreciable rate (such as 60 words/min) over an automatic teletype system, that the value of $C/N_{1\text{ kc}}$ may be approximately 18 db for 0.1 percent errors [15]. The very great increase in required powers for communication is readily seen.

[2] In order to conform with general usage, it should be noted that as the ability to accurately compare frequency increases, the value of ϵ actually decreases.

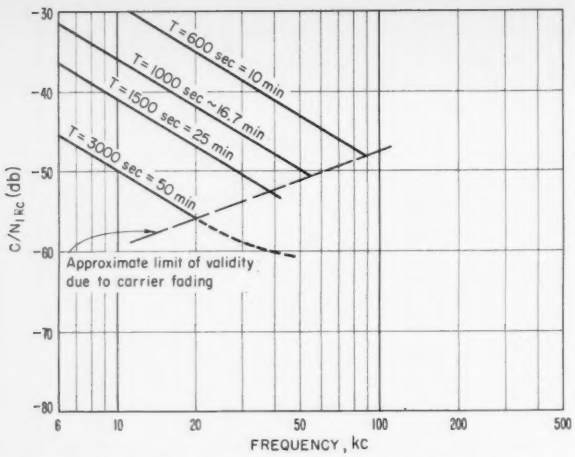

T = 600 sec = 10 min
T = 1000 sec ~16.7 min
T = 1500 sec = 25 min
T = 3000 sec = 50 min
Approximate limit of validity due to carrier fading
C/N1 kc (db)
FREQUENCY, KC

FIGURE 4. *Carrier-to-noise requirements as a function of frequency for a frequency comparison with a precision of 1 part in 10^9 in observing periods indicated, based on eq (C6).*

Eq (C6):

$$C/N_{1\,kc}(d\ell) = -106 - 20 \log \epsilon - 20 \log f_{kc} - 30 \log T,$$

where ϵ = required precision
f_{kc} = frequency in kc
T = observing period (seconds).
NOTE.—Equivalent receiver bandwidth $\approx 1/T$.

5. Calculations of Required Radiated Power

The next term in eq (1) is T_z, the allowance which must be made to provide the level of performance specified for a given percentage of all hours. A value of 90 percent of all hours was chosen as being reasonable for the type of service involved. T_{zn}, shown as the lower curve in figure 5, represents the additional allowance which must be made for the variability in atmospheric noise levels, and T_{zs} is the allowance for signal level variability. T_z is obtained by taking the square root of the sum of the squares of the individual variances of T_{zn} and T_{zs}. This yields the upper curve, and it can be noted that both the noise level and anticipated transmission loss variability combine to require a margin that increases with frequency above 20 kc. The inserted curve on figure 5 may be used to determine T_z for less than 90 percent of all hours. For example, where T_z for 90 percent of all hours is 10 db at 20 kc, T_z for 74 percent of all hours is 50 percent of T_z (90) or 5 db at 20 kc.

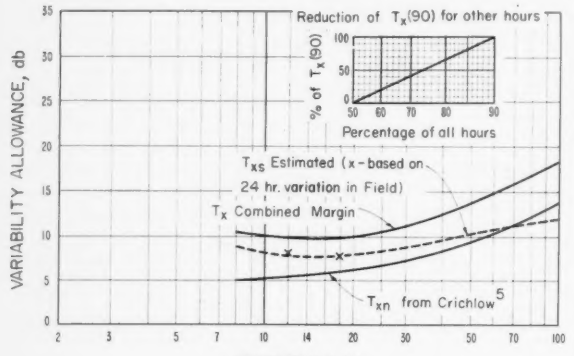

Reduction of Tx(90) for other hours
% of Tx(90)
Percentage of all hours
Txs Estimated (x-based on 24 hr. variation in Field)
Tx Combined Margin
Txn from Crichlow5
VARIABILITY ALLOWANCE, db
FREQUENCY, KC

FIGURE 5. *Signal fading and noise variation margin for 90 percent of all hours.*

When all of these factors are properly combined, we obtain the curve shown in figure 6 which indicates minimum required radiated power at around 19 kc. It is interesting to observe that this low-loss vlf pass band is very close to the region where the majority of vlf transmitters are presently located.

The 2 kw of radiated power required for the 50-min observing period must be considered as a minimum value. It may be desirable to increase this value of radiated power to provide a margin of performance, and before a final choice is made, estimates should be prepared of the initial and operating costs versus radiated power for stations with capabilities ranging from 1 kw to 200 kw.

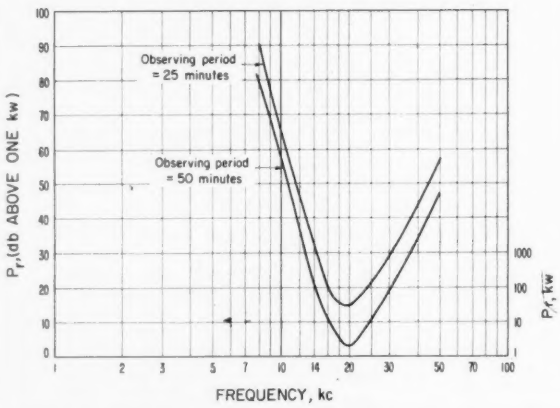


FIGURE 6. *Radiated power required for a worldwide standard frequency broadcasting station.*

Calculations based on a primarily daylight path of 17,000 km to the vicinity of Java or Madagascar (estimated most difficult service areas). Transmitter located near 40° N, 105° W.

6. Antenna Considerations

The factors involved in the design of an antenna to radiate economically the powers indicated in figure 6 are extremely complex. To give some idea of the effect of antenna characteristics on the choice of an optimum frequency, we can consider the transmitter power requirements based on the use of an antenna with properties similar to the NSS Annapolis, Md., antenna. The characteristics of this antenna expressed as antenna loss in decibels as a function of frequency are shown in figure 7. Although this is a rather complex antenna, consisting of nine 600-ft towers and capable of radiating approximately 200 kw at 20 kc, it can be shown that very similar loss characteristics could be obtained with a less expensive antenna if the radiated-power requirements were lowered.

Assuming the antenna characteristics in figure 7, we have shown by the curves in figure 8 the transmitter power required to furnish the service specified

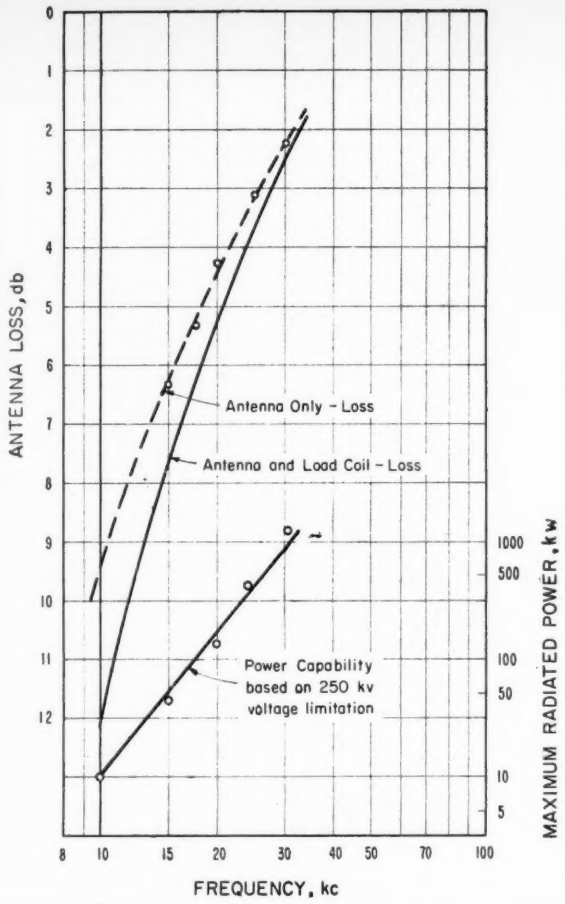

ANTENNA LOSS, db
MAXIMUM RADIATED POWER, kw
FREQUENCY, kc
Antenna Only - Loss
Antenna and Load Coil - Loss
Power Capability based on 250 kv voltage limitation

FIGURE 7. *Antenna power loss and maximum power capability.*

NSS Annapolis: nine 600-ft towers.

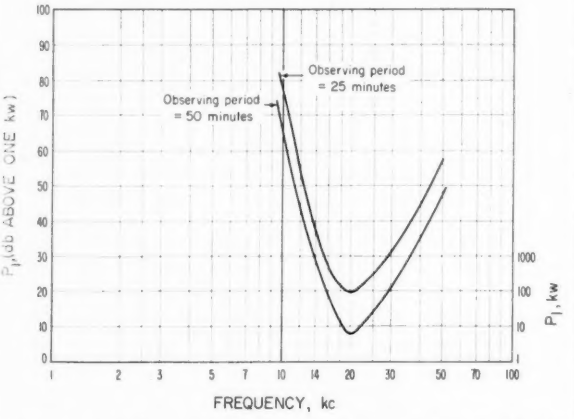


FIGURE 8. *Antenna input power requirements for a worldwide standard frequency broadcasting station.*

Calculations based on a primarily daylight path of 17,000 km to the vicinity of Java or Madagascar (estimated most difficult service areas). Transmitter located near 40° N, 105° W.

as a function of frequency. It should be emphasized that even if it were feasible to generate the powers which appear to be required at the lower frequencies, around 10 kc, that existing vlf antennas are not capable of radiating such powers due to corona limitations. This fact would be true of all known or planned antennas.

This voltage limiting effect which makes it more expensive to radiate a given amount of power as the frequency is decreased in this range must also be considered before a final choice of optimum frequency is made. It is not likely that this factor will increase the optimum frequency very much, but its influence should not be overlooked in a detailed study.

The authors are indebted to the following: K. A. Norton, J. R. Wait, W. W. Brown, F. M. Malone, A. H. Morgan, L. R. O. Storey, and W. D. George for helpful discussions and information; and to Mrs. W. Mau for her assistance in the preparation of the manuscript.

7. Appendixes

7.1. Appendix A: VLF Vertical Electric Field Intensity

Wait [7] has shown that the vertical electric field from a vertical current source can be written as

$$E_r = E_o \left[\frac{d/a}{\sin(d/a)} \right]^{1/2} \frac{(d/\lambda)^{1/2}}{h/\lambda} A/U, \qquad \text{(A1)}$$

where E_o is the field of the source on a perfectly conducting flat earth at a distance d from the source, d is this distance, a is the earth's radius, h is the height to the ionosphere, and λ is the wavelength, all in the same units. A/U represents a more complex expression in the original formula which in essence is the ratio between the actual field and the unattenuated field that would result if both earth and ionosphere were perfectly conducting spherical shells.

Using the well known relations

$$E_o = \frac{3\times 10^5 \sqrt{P_r\,(\text{kw})}}{d\,(\text{km})}\ \mu\text{v/m}, \qquad \text{(A2)}$$

and

$$\lambda\,(\text{km}) = \frac{300}{f\,(\text{kc})}, \qquad \text{(A3)}$$

we can write

$$E_r = \frac{5.2\times 10^6 \sqrt{P_r\,(\text{kw})}}{[a\ \sin(d/a)]^{1/2} \cdot \sqrt{f\,(\text{kc})} \cdot h\,(\text{km})} \cdot A/U. \qquad \text{(A4)}$$

where it is interesting to observe that if A/U remained constant, E_r would be inversely related to

h and $\sqrt{f}$. In the actual case A/U is a function of both of these factors. The complete expression for A/U involves the summation of many modes; however, at a distance in excess of 2,000 km, the first order mode is dominant and the attenuation can be expressed in terms of a simple exponential decay usually given in terms of α (decibels per 1,000 km). When this is employed and (A4) is expressed in logarithmic terms, we obtain eq (2).

7.2. Appendix B: Transmission-Path Phase Stability and Signal-Integration Times Required for Specified Frequency-Comparison Precisions

When a standard frequency broadcast is propagated via the ionosphere, Pierce [2] has shown that the diurnal changes of ionosphere heights, as well as the apparent roughness of the ionosphere, introduce effective transmission time variations. These variations in propagation time introduce phase variations on the received carrier, and the differential of these phase variations with respect to time represent the apparent instantaneous frequency deviations of the received carrier from the standard broadcast frequency.

The present state of the art in oscillator development has reached a point where these instantaneous frequency deviations introduced by propagation are considerably greater than the instabilities of the oscillator itself. Therefore, it becomes necessary to devise methods of reducing the effect of these propagation path induced frequency variations when accurate comparisons of frequency are desired. It has been shown by Pierce [2] that increasing the time of observation permits a more precise frequency comparison, and it is the purpose of this analysis to describe some of the factors involved in the amount of precision which is obtainable in a given length of time.

We shall first assume a very stable oscillator at the transmitting location with a frequency f_s. At the receiving location the frequency received f_r is not necessarily equal to f_s since the propagation path introduces a delay which varies with time. This variation in delay can be separated into a large diurnal variation with a smaller random component [2]. The standard deviation of this latter component expressed in degrees is defined as σ_Ω where typical values from Norton [16] are shown in figure 9 as a function of frequency for a single-hop path.

Assuming a condition where the diurnal phase is constant, the output of a phase detector at the receiving location relative to a highly stable local frequency source, will have the form shown in figure 10. The average observed frequency difference between the local oscillator f_l and the received carrier f_r will depend on the type of observation made.

Case I: The instantaneous frequency as derived by Carson [21] is the time derivative of phase, and the instantaneous frequency difference between f_l and f_r is

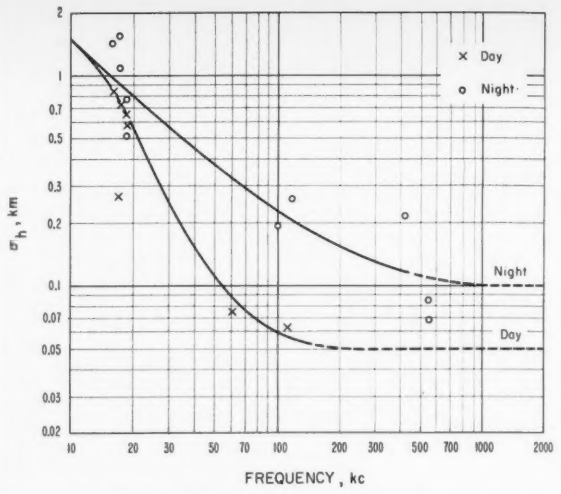


FIGURE 9. *Median effective ionospheric roughness parameter, σ_h, obtained from observations of σ_Ω, σ_Ω (degrees) $\equiv 2.4\ f_{kc}\ cos\ \phi_i\ \sigma_h$ (km).*

(See table B1 for sources)

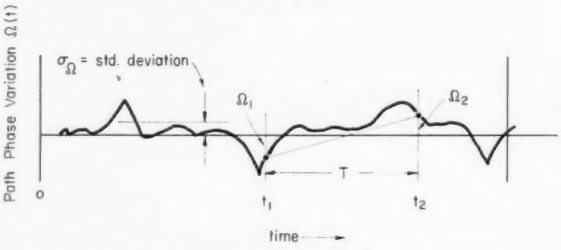


FIGURE 10. *Transmission-path phase variation.*

$$f_d\ (\text{instantaneous}) = \frac{1}{360} \cdot \frac{d}{dt}\ \Omega(t), \qquad \text{(B1)}$$

where f_d is in cps and $\Omega\ (t)$ is the phase difference in degrees between the carrier field at the receiver and a constant local reference. Measurements employing an instantaneous comparison of f_l and f_r will obviously have large excursions and, in view of this, we can consider averaging over a given interval T as shown in figure 10.

Case II: If the receiving system only sampled $\Omega\ (t)$ at times t_1 and t_2, we would obtain

$$f_d\ (\text{avg}) = \frac{\Omega_2 - \Omega_1}{360\ T}. \qquad \text{(B2)}$$

When T is short compared to the average carrier amplitude fade or phase variation period, τ, $f_d(\text{avg}) = f_d\ (\text{instantaneous})$; while if T is long compared to a fade, Ω_1 and Ω_2 are independent. If the standard deviation of the received carrier phase σ_Ω is known, we obtain a total phase variance of $\sqrt{2}\,\sigma_\Omega$ and as a result

$$\sigma f_d\ (\text{avg}) = \frac{\sqrt{2}\sigma_\Omega}{360\ T}, \quad T > \tau \qquad \text{(B3)}$$

where σf_d (avg) is the standard deviation of the average frequency difference over a period T from the true transmitted frequency.

Case III: A more desirable type of observation would be to integrate $\Omega(t)$ over each consecutive period T such that the standard deviation of Ω out of the integrator at the end of each period is [22]

$$\frac{\sigma_o}{\sigma_i} \approx \frac{0.7}{\sqrt{T/\tau}}, \quad T/\tau > 1 \tag{B4}$$

where τ is the average fade period of the received carrier. Using this type of observation and combining (B3) and (B4) we obtain

$$\sigma f_d \text{ (integrated avg)} \approx \frac{\sigma_\Omega \tau^{1/2}}{360\ T^{3/2}}, \quad T/\tau >> 1. \tag{B5}$$

This standard deviation of frequency difference can be used to obtain the precision of frequency comparison, ϵ, where

$$\epsilon = \sigma f_d / f_t. \tag{B6}$$

In all cases, we will consider the phase for a single ray path and assume that the phase variation will increase directly with the square root of the number of ionospheric reflections m. It is also assumed that these paths are near grazing incidence which produces a minimum value for cos φ_i of 0.15 where φ_i is the angle of ionospheric incidence. Employing these assumptions we obtain from figure 9

$$\sigma_\Omega \text{ (deg)} = 360 \times 10^{-6} \sqrt{m} \cdot f \cdot \sigma_h. \tag{B7}$$

Since in eq (B1) σ_Ω is not given, we can observe that for phase variations of the type considered

$$\sigma f_d = \frac{k \sigma_\Omega}{360 \tau} \tag{B8}$$

where τ is the average fade or phase variation period and experience has indicated values for k in the order of 4.

Combining eq (B6) and (B7) with (B8), (B3), and (B5), respectively, we obtain the following precisions for each of the three cases considered.

Case I—instantaneous precision

$$\epsilon_1 \approx \frac{k\sqrt{m} \cdot \sigma_h}{10^6 \tau}, \quad (T < 0.1\tau). \tag{B9}$$

Case II—average precision

$$\epsilon_2 = \frac{\sqrt{2} \cdot \sqrt{m} \cdot \sigma_h}{10^6 T}, \quad (T > 3\tau). \tag{B10}$$

Case III—integrated average precision

$$\epsilon_3 = \frac{\tau^{1/2} \cdot \sqrt{m} \cdot \sigma_h}{10^6 T^{3/2}}, \quad (T > 3\tau). \tag{B11}$$

Since the average fade duration τ is required in cases I and III, an estimate of τ over the frequency range in question is given in figure 11. Combining these values with those obtained from figure 9, we obtain the estimated precision obtainable over a 17,000-km path at 20 and 50 kc, neglecting noise, shown in figure 12. The solid curves are experimental values from Pierce [2] for a somewhat shorter path. The single ray path analysis in effect implies that ϵ will increase with the square root of path distance. Actual data on the distance effect is unfortunately not available; however, it is expected that in the 20-kc vlf region and for long paths several ray paths will combine so as to reduce the fluctuations in phase. This will reduce ϵ below the values indicated for the single ray path.

It is interesting to observe that in the 16- to 20-kc region and for large values of T, ϵ is proportional to $1/T$. This probably results since the periods required are so long that the path phase is not independent of diurnal phase effects and eq (B11) does not apply. In the 50- to 60-kc region T is greater than 3τ for periods less than 20 min. As a result eq (B11) can apply. From Pierce's observations it is also seen that for large values of T, ϵ is now proportional to $1/T^{3/2}$.

TABLE B–1. *Estimates of the variance of phase on ionospheric paths*

No.	Frequency	Distance	Source	Time	Month	$\sigma_{obs.}$	S	$\sqrt{2(1-\rho)^2}$	σ_Ω	Mode	cos ϕ	σ_h
	kc	*Statute miles*				*Degrees*	λ			*m*		*km*$_i$
1	16	3,230	Pierce [2]	Day	February	8.06			8.06	3	0.146	0.829
2	16	3,230	do	Night	do	16.1			16.1	3	.170	1.424
3	17.2	640	Redgment [17]	Day	September	1.9	10	1.00	1.9	1	.174	0.265
4	17.2	640	do	Night	do	9.3	10	1.00	9.3	1	.206	1.093
5	17.2	640	do	Day	January	5.2	10	1.00	5.2	1	.174	0.726
6	17.2	640	do	Night	do	13.3	10	1.00	13.3	1	.206	1.563
7	18.4	3,488	do	Day	June and July	7.1	10	1.00	7.1	3	.144	0.643
8	18.4	3,488	do	Night	do	10.0	10	1.00	10.0	3	.169	.775
9	18.4	3,488	do	Day	December	6.5	10	1.00	6.5	3	.144	.588
10	18.4	3,488	do	Night	do	6.5	10	1.00	6.5	3	.169	.503
11	60	3,230	Pierce [2]	Day	do	2.8			2.8	3	.146	.077
12	100	2,350	Doherty [18]	Night	April	10.8	>40	1.41	7.64	1	.165	.193
13	115	780	Florman [19]	do	March	13.3	11	1.00	13.3	1	.189	.255
14	115	780	do	Day	do	2.7	11	1.00	2.7	1	.158	.062
15	418	1,250	do	Night	do	51	41	1.41	36.1	1	.167	.216
16	543	380	Brennan [20]	do	January to April	34	3.39	0.83	41.0	1	.371	.085
17	556	454	Redgment [17]	do	April and May	19	1.85	0.63	30.2	1	.334	.068

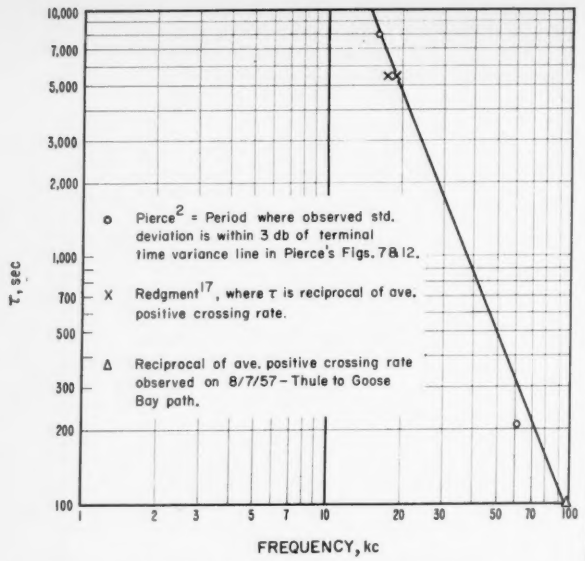


FIGURE 11. *Estimated fade period τ for a long distance daytime path* d>10,000 *km.*

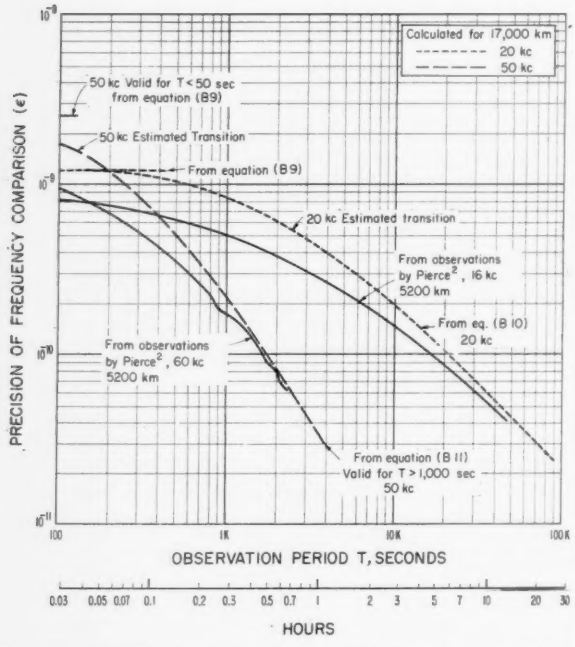


FIGURE 12. *Calculated precision of frequency comparison obtainable for a* 17,000-*km daytime path.*

Theoretical curves based on ionospheric roughness given in figure 9, estimated fade period τ given in figure 11, and an assumed $m=7$ ionospheric reflections.

7.3. Appendix C: Carrier-to-Noise Requirements for Precise Frequency Measurement

In all physical measurements there is a noise level which limits the degree of precision with which a given measurement can be made. The precision of frequency measurement can be determined from the precision with which a given number of degrees of carrier phase rotation can be measured. For example, if a carrier is contaminated with a small amount of narrow-band noise, the maximum phase variation of the resultant is $\phi_{max}=n/c$ where n and c are peak noise and carrier values. If we now desire the standard deviation of phase σ_ϕ, we can replace n with N and c with $\sqrt{2}\ C$ and obtain

$$\sigma_\phi \approx \frac{1}{\sqrt{2}}\frac{N}{C}, \tag{C1}$$

where C and N are the rms carrier and noise voltages [24].

If we now assume that our true period is $2\pi r$ radians where r is the number of whole cycles in our observing period and since ϕ is normally distributed, the total uncertainty or phase jitter is $\sqrt{2}\ \sigma_\phi$ radians, and the precision is given by

$$\epsilon=\sqrt{2}\sigma_\phi/2\pi r,$$

or we can express the precision obtainable as

$$\epsilon \sim \frac{N/C}{2\pi r}, \quad C/N \gg 1. \tag{C2}$$

If we now consider a receiving system where we integrate the phase output over an observing period, T, and compare it with the integrated value of the preceding period as in eq (B4), we obtain a reduction in phase jitter of approximately $(T\cdot BW)^{-1/2}$ as long as $T>1/BW$, where BW is the effective rf bandwith of the receiver before the final integration. Employing this correction to (C2), and observing that $r=T\cdot f$, where T is the period of measurement in seconds and f is the frequency in cycles per second, we obtain

$$\epsilon \approx \frac{N/C}{2\pi f T^{3/2} BW^{1/2}}. \tag{C3}$$

It should be emphasized that this expression applies only to the type of observation described, and it is not necessarily equal to the optimum obtainable.

Before employing (C3) we must be sure that inaccuracies of measurement which can occur in some systems do not arise due to a miscount of a whole cycle. This requirement is met by being sure that the noise envelope is less than the carrier for a specified percent of all time. Since the noise envelope is Rayleigh distributed, a protection of 1,000 to 1 is obtained with $C/N=8.5$ db, and 10,000 to 1 with 9.6 db. The actual probability of having the noise envelope remain below the carrier envelope for a

whole period T is found by taking the probability of one independent sample being correct and raising it to a power determined by the total number of independent samples which could be made in the whole sample. This power is approximately equal to $BW \cdot T$ which for a BW of 0.01 cps, and a T of 1 hr is 36. If we choose the $C/N = 8.5$ db case, we must raise 0.999 to the 36th power which yields a 96.7 percent probability of being free from errors of a whole cycle or more. It should be pointed out that this requirement which in effect places a threshold on the C/N in the receiver pass band is not always necessary since it is possible to locally introduce the correct number of whole cycles if the local standard is approximately correct.

Returning to (C3) and expressing C/N in decibels, we can write

$$C/N \text{ (db)} \approx -76 - 20 \log \epsilon - 20 \log f(\text{kc}) - 30 \log T - 10 \log BW. \quad \text{(C4)}$$

Since many different receiver combinations can be employed, it is desirable to obtain our results in terms of the rms carrier to rms noise in a 1-kc bandwidth expressed in decibels which is obtainable from

$$C/N_{1\text{kc}} = C/N \text{ (db)} - 10 \log \frac{1{,}000}{BW}, \quad \text{(C5)}$$

where BW is the receiver bandwidth in cycles per second. Combining (C4) and (C5) we obtain

$$C/N_{1\text{kc}} \approx -106 - 20 \log \epsilon - 20 \log f(\text{kc}) - 30 \log T, \quad \text{(C6)}$$

where it is seen that for a coherent type of detection the results are independent of receiver IF bandwidth. Typical ranges of $C/N_{1\text{kc}}$ for a precision of frequency comparison of one part in 10^9 are shown as a function of T in figure 13, and as a function of f in figure 4.

When T approaches the carrier fade period τ, described in appendix B, the effective carrier amplitude is reduced and (C6) no longer holds. In fact when T is several times τ, it is expected that the last term in (C6) will approach a decrease of 10 db per decade rather than 30. Since this effect is quite complex, including the shape of the power spectrum of the received field, we have only shown estimated $C/N_{1\text{kc}}$ values by dashed lines in this region. In addition, diurnal effects are expected to limit the validity of all these curves to values less than 15-K sec.

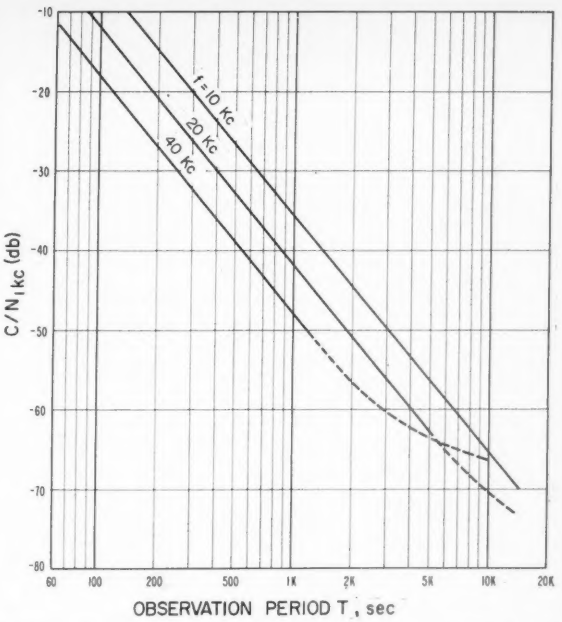


FIGURE 13. *Calculated carrier-to-noise requirements for a frequency comparison with a precision ϵ of 1 part in 10^9 in a period of T seconds.*

Calculations based on eq (C6):
$C/N_{1\text{kc}}(db) = -106 - 20 \log \epsilon - 20 \log f_{\text{kc}} - 30 \log T$
where: ϵ = required precision
f_{kc} = frequency in kc
T = observing period (seconds).

8. References

[1] J. A. Pierce, H. T. Mitchel, and L. Essen, World-wide frequency and time comparisons by means of radio transmissions, Nature **174,** 922 (1954).

[2] J. A. Pierce, Intercontinental frequency comparison by vlf radio transmission, Proc. IRE **45,** 794 (1957).

[3] A. H. Allan, D. D. Crombie, and W. A. Penton, Frequency variations in New Zealand of 16 kc/s transmissions from GBR Rugby, Nature **177,** 178 (1956).

[4] W. D. George, NBS Boulder Laboratories. Informal communication.

[5] W. Q. Crichlow, D. F. Smith, R. N. Morton, and W. R. Corliss, Worldwide radio noise levels expected in the frequency band 10 kilocycles to 100 megacycles, NBS Circ. 557 (1955).

[6] Proc. IRE **45,** 6 (1957), vlf issue.

[7] J. R. Wait, The mode theory of vlf ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (1957).

[8] H. J. Round, T. L. Eckersley, K. Tremellen and F. C. Lunnon, Report on measurements made on signal strength at great distances during 1922 and 1923 by an expedition sent to Australia, J. Inst. Elec. Engrs. (London) **63,** 933 (1925).

[9] J. A. Pierce, Sky wave field intensity I. Low and very low frequencies, Cruft Laboratory, Harvard U., Technical Report 158, Sept. 1, 1952.

[10] J. L. Heritage and S. Weisbrod, A study of signal versus distance data at vlf, Symposium on the propagation of vlf radio waves, Boulder, Colo., Jan. 1957, Vol. III.

[11] J. R. Wait, An extension to the mode theory of vlf ionospheric propagation, J. Geophys. Research **63,** 125 (1958).

[12] A. D. Watt, and E. L. Maxwell, Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (1957).

[13] W. L. Taylor, and L. J. Lange, Some Characteristics of vlf propagation using atmospheric waveforms, Proceedings of the second conference on Atmospheric Electricity at Portsmouth, N.H., May, 1958 (Pergamon Press).

[14] A. D. Watt, and E. L. Maxwell, Measured statistical characteristics of vlf atmospheric radio noise, Proc. IRE **45,** 55 (1957).

[15] A. D. Watt, R. M. Coon, E. L. Maxwell, and R. W. Plush, Performance of some radio systems in the presence of thermal and atmospheric noise, Proc. IRE **46,** 1914 (1958).
[16] K. A. Norton, Transmission loss in radio propagation II, to be published.
[17] P. G. Redgment and D. W. Watson, Phase-correlation of medium and very-low-frequency waves using a baseline of several wavelengths, Admiralty Signal and Radar Est., Lythe Hill House, Haslemere, Surrey, England, Monograph No. 836, October, 1948.
[18] R. H. Doherty, NBS Boulder Laboratories. Informal communication.
[19] E. F. Florman, NBS Boulder Laboratories. Informal communication.
[20] D. G. Brennan and M. Lindeman Phillips, Phase and amplitude variability in medium-frequency ionospheric transmission, Technical Report No. 93, Massachusetts Institute of Technology Lincoln Laboratory, September (1957).
[21] J. R. Carson, Notes on the theory of modulation, Proc. IRE **10,** 57 (1922).
[22] A. D. Watt, Statistical characteristics of sampled and integrated A-M and F-M noise, NRL Report 3856 (Oct. 22, 1951).
[23] R. A. Helliwell and E. Gehrels, Observations of magnetoionic duct propagation using man-made signals of very low frequency, Proc. IRE **46,** 785 (1958).
[24] K. A. Norton, E. L. Shultz, and Helen Yarbrough, The probability distribution of the phase of the resultant vector sum of a constant vector plus a Rayleigh distributed vector, J. Appl. Phys. **23,** 137 (1952).
[25] J. R. Wait, A study of vlf field strength data—both old and new, Geofisica pura e applicata **41** 73 (1958/III).

BOULDER, COLO. (Paper 63D1–6)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Measurements of Phase Stability Over a Low-Level Tropospheric Path

M. C. Thompson, Jr., and H. B. Janes

(February 19, 1959)

A knowledge of the statistics of atmosphere-induced variations in the phase of the received signal (i.e., variations in electrical path length) is essential in evaluating the reliability of any system using radio waves for measuring distance and/or velocity. This paper describes an analysis of phase variations measured at 9,400 megacycles per second over a 9.4-mile path near Boulder, Colorado, during a 40-hour period in September 1958. The power spectral density of these variations is shown to be approximately proportional to $f^{-2.3}$ over a wide range of frequencies throughout the period of recording. The long-term phase variations are closely correlated with atmospheric refractivity measurements made at the path terminals.

1. Introduction

The National Bureau of Standards has been conducting a series of experiments to study the instability of the phase of vhf, uhf, and microwave radio signals transmitted over line-of-sight paths.[1] These experiments have included measurements of phase stability over "slanted" paths such as the Pikes Peak to Garden of the Gods path near Colorado Springs, Colo., and the Haleakala summit to Puunene path on the island of Maui, T.H. This paper describes some measurements of phase stability made at 9,400 Mc over a 9.4-mile path from Green Mountain mesa to Table mesa near Boulder, Colo. This path differs from the ones mentioned above in that both terminals are at approximately the same altitude. The data consist of: (1) 40 hr of continuous recording of long-term variations in the phase of the radio signal; (2) 40 hr of continuous atmospheric temperature, pressure and relative humidity recordings; and (3) 21 samples (each approximately 5 min long) of short-term phase variations. The analysis of these data consists of: (1) Estimation of the power density spectrum of the phase variations, using both long-term and short-term recordings; (2) estimation of the power density spectrum of the corresponding variations in radiofrequency of the received signal; and (3) correlation of the long-term variations in electrical path length (from the phase records) and surface refractivity (from the meteorological records).

2. Path Terrain and Weather

Figure 1 shows the location of the propagation path. It extends from Green Mountain mesa about ½ mile west of the NBS, Boulder Laboratories, to Table mesa about 9 miles north of the Boulder Laboratories. The path length is 9.4 miles. As shown in figure 2, each antenna was located at the edge of a steep slope to avoid ground reflections in the immediate vicinity. The path passed over two valleys which slope downward from west to east and are separated by a relatively flat mesa. This mesa contains a small lake (Roberts Lake), and appears to constitute the major ground-reflection area.

The measurements were made during the period 1100, September 15, to 1625, September 17, 1958.

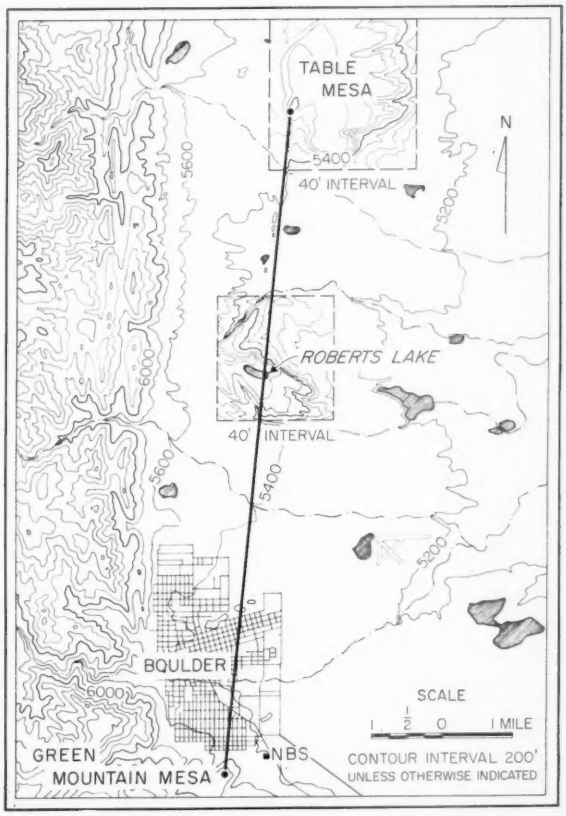


FIGURE 1. *Green Mountain mesa—Table mesa path.*

[1] J. W. Herbstreit and M. C. Thompson, Jr., Measurements of the phase of radio waves received over transmission paths with electrical lengths varying as a result of atmospheric turbulence, Proc. IRE **43**, 1391 (1955).

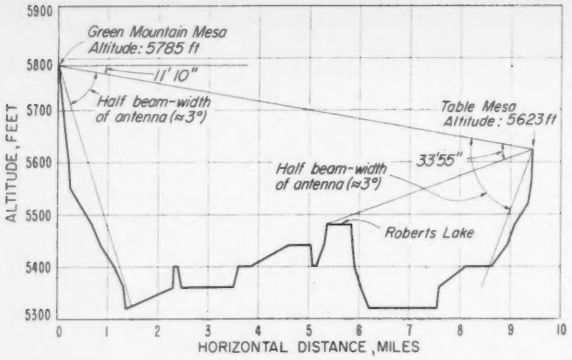


FIGURE 2. *Terrain profile of Green Mountain mesa—Table mesa path.*

Weather conditions over the path during this period are summarized graphically in figures 3 and 4, which show the wind velocity, barometric pressure, temperature, and relative humidity observed at the path terminals, and in the notes on figure 5. A cold front moved across the path in a direction approximately normal to the path during the period 1000 to 1200 on September 15. The winds, temperature drop, and increase in relative humidity are indicated by figures 3 and 4. From about 1300 until 1830 the sky was overcast in all directions and a light rain fell steadily over all or part of the path, amounting to about .17 in. total accumulation. During the night patches of ground fog developed. As daylight came on September 16, from the tops of the mesas (the path terminals) fog could be seen for many miles to the northeast, east, and southeast, lying in the low areas of the terrain. Following sunrise at about 0540, these pockets of fog gradually disappeared until by 1000 the entire area was clear, with excellent visibility and partially cloudy sky. Pikes Peak, at a distance of 90 miles, was clearly visible from the Table mesa site from the onset of daylight. The remainder of the run was conducted in completely clear weather conditions.

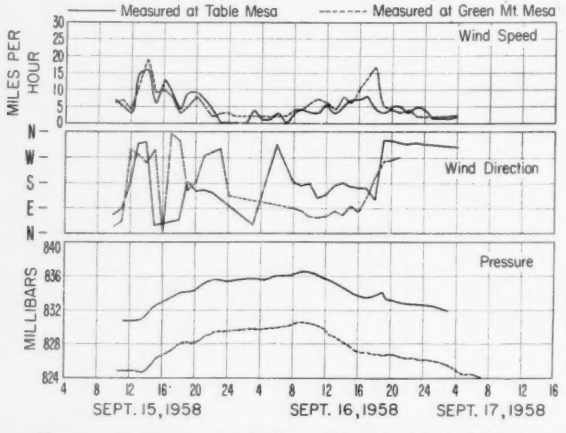


FIGURE 3. *Summary of phase and meteorological data.*

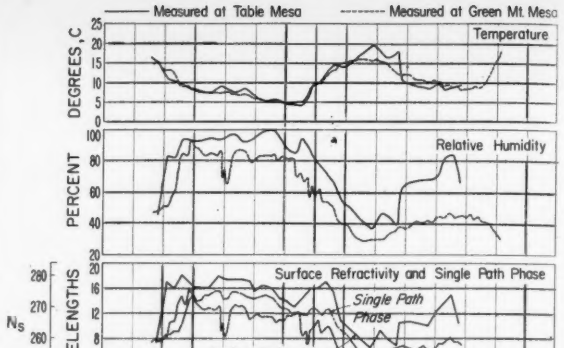


FIGURE 4. *Summary of phase and meteorological data.*

3. Phase Measurement System

The phase variations were measured by a system described earlier.[2] It consists essentially of a quasi-symmetrical arrangement in which two signals are compared after having traversed the same path in opposite directions. The measuring system employed had a noise level of the order of 0.1 deg of phase due to combined thermal noise and electronic instabilities. At the operating wave length of about 1 in., this uncertainty corresponds to a physical displacement of either antenna (along the line of the path) of about 350 μ in. Thus, the need for rigid antenna supports is apparent, and solid concrete piers 2 ft in diam were used as mounts for the 18 in. parabolic dishes. The antennas were from 2 to 5 ft above the immediate ground level. The frequency of the microwave measuring signal was stabilized to a few parts in 10^8 using a phase-lock technique.[2]

The actual phase comparison was performed by a commercial electronic phase meter in conjunction with an electromechanical instrument developed at the Boulder Laboratories.

Two recording circuits were used to measure the phase variations in order to study the power spectrum over as wide a range of frequencies as possible. The output of one, referred to as the low-pass recording, consisted of a Varian chart moving at the rate of 0.1 ipm. A full-scale sweep of the pen represented a change of 1,800 deg in the difference between the phase of the signals received at the two-path terminals (almost) simultaneously. The "single-path phase" data discussed here was obtained by simply dividing this "two-way phase" by two. The frequency response characteristic of this recording circuit is essentially flat up to about 1 cps. However, since the primary purpose of this record was to measure gross changes in phase, the dynamic range and chart speed used necessarily precluded the study of spectral frequencies above about 1 cpm. It was

[2] M. C. Thompson, Jr., M. J. Vetter, Single path phase measuring system for three-centimeter radio waves, Rev. Sci. Instr. (1959).

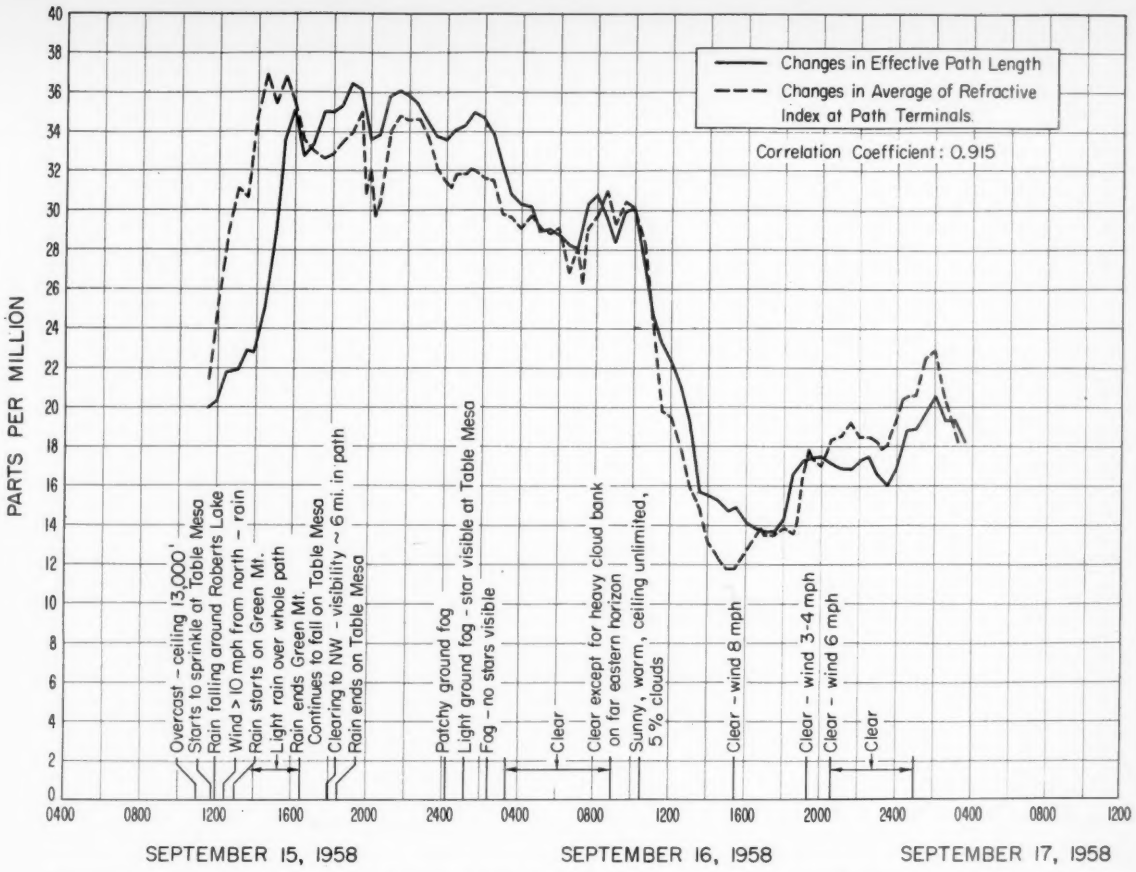

Changes in Effective Path Length
Changes in Average of Refractive Index at Path Terminals.
Correlation Coefficient: 0.915
PARTS PER MILLION
Overcast – ceiling 13,000'
Starts to sprinkle at Table Mesa
Rain falling around Roberts Lake
Wind > 10 mph from north – rain
Rain starts on Green Mt.
Light rain over whole path
Rain ends Green Mt.
Continues to fall on Table Mesa
Clearing to NW - visibility ~ 6 mi. in path
Rain ends on Table Mesa
Patchy ground fog
Light ground fog - star visible at Table Mesa
Fog – no stars visible
Clear
Clear except for heavy cloud bank on far eastern horizon
Sunny, warm, ceiling unlimited, 5% clouds
Clear – wind 8 mph
Clear – wind 3-4 mph
Clear – wind 6 mph
Clear
SEPTEMBER 15, 1958
SEPTEMBER 16, 1958
SEPTEMBER 17, 1958

FIGURE 5. *Comparison of variations in effective path length with average of surface refractive index at path terminals.*

possible to record a wide range of values without loss of continuity because of an automatic scale-shifting arrangement which prevented the trace from going off scale.

The output of the second phase recording circuit, referred to as the band-pass recording, consisted of a Sanborn chart having full-scale ranges of 1.6 to 8 single-path deg and moving at the rate of about 60 ipm. Its frequency response characteristic consisted of a pass-band extending from about 0.03 cps to about 30 cps. This circuit was operated for 21 periods, each about 5 min in length, distributed over the 40-hr period.

4. Power Spectrum of Phase Variations

One of the principal objectives of the experiment was to obtain information on the power spectrum of time variations in the phase of a microwave signal propagated over a line-of-sight path. The power spectrum might better be considered as the variance density spectrum, since it shows graphically how much of the total variance (mean squared deviations from the sample mean) was contributed by variations occurring in each of several narrow frequency bands distributed over the pass-band of the recording equipment. The method of estimating the power spectrum used here is that given by Blackman and Tukey.[3] Briefly, it consists of computing the Fourier cosine transform of the autocovariance function and applying appropriate weighting functions to the result.

Seven of the twenty-one Sanborn recording periods were chosen for the power spectrum analysis, the only criterion used in their selection was that they be fairly evenly spaced over the 40-hr period. The analysis of each period consisted of first computing the spectrum in the range 0.83 to 12.5 cps for two adjacent data segments each 0.5 min in length, using a sampling interval of 0.04 sec. Then the two segments were combined, the sampling interval increased to 0.2 sec and the spectrum computed for the range 0.083 to 2.5 cps. This yielded 14 sample spectra in the higher frequency range and 7 in the lower range of the Sanborn records.

[3] R. B. Blackman and J. W. Tukey, The measurement of power spectra from the point of view of communication engineering, Bell System Tech. J., Pts. 1, 2, **37**, 185, 485 (1958).

From consideration of the frequency response of the recording circuit, it is felt that the seven lower range spectra are not seriously contaminated by frequencies below 0.083 cps. However, the data for the 14 higher range spectra were subjected to a digital trend removal process,[4] since they would otherwise be contaminated by variance contributed at frequencies between the lowest spectral frequency estimated (0.83 cps) and the lower limit of the pass-band (about 0.03 cps). (In all cases, the spectra were corrected to allow for the recording circuit response.)

[4] This process consists simply of computing the deviations of the variable from a running average, computing the power spectrum of these deviations, and correcting the result for the frequency response of this high-pass "filter." This, in effect, filters out variations at spectral frequencies below a cutoff frequency which is determined by the averaging time.

The analysis of the Varian low-pass data consisted of first digitizing the entire 40-hr run with a sampling interval of 0.5 min. Then five sample spectra were computed, each using 6 to 8 hr of data, covering the range from 0.00111 to 0.0167 cps, ($\frac{1}{15}$ to 1 cpm). Then two sample spectra were computed, each using 15 hr of data and covering the range from 0.000278 to 0.00417 cps ($\frac{1}{60}$ to $\frac{1}{4}$ cpm), and finally, one spectrum utilizing all 40 hr of data and covering the frequencies from 0.0000926 to 0.00139 cps (8 cycles per day to $\frac{1}{12}$ cpm). In each case, the data were subjected to the trend removal process to minimize the effects of variations at frequencies below the range covered in the computation.

The individual spectral densities for each sample are plotted in figure 6. The points shown as open

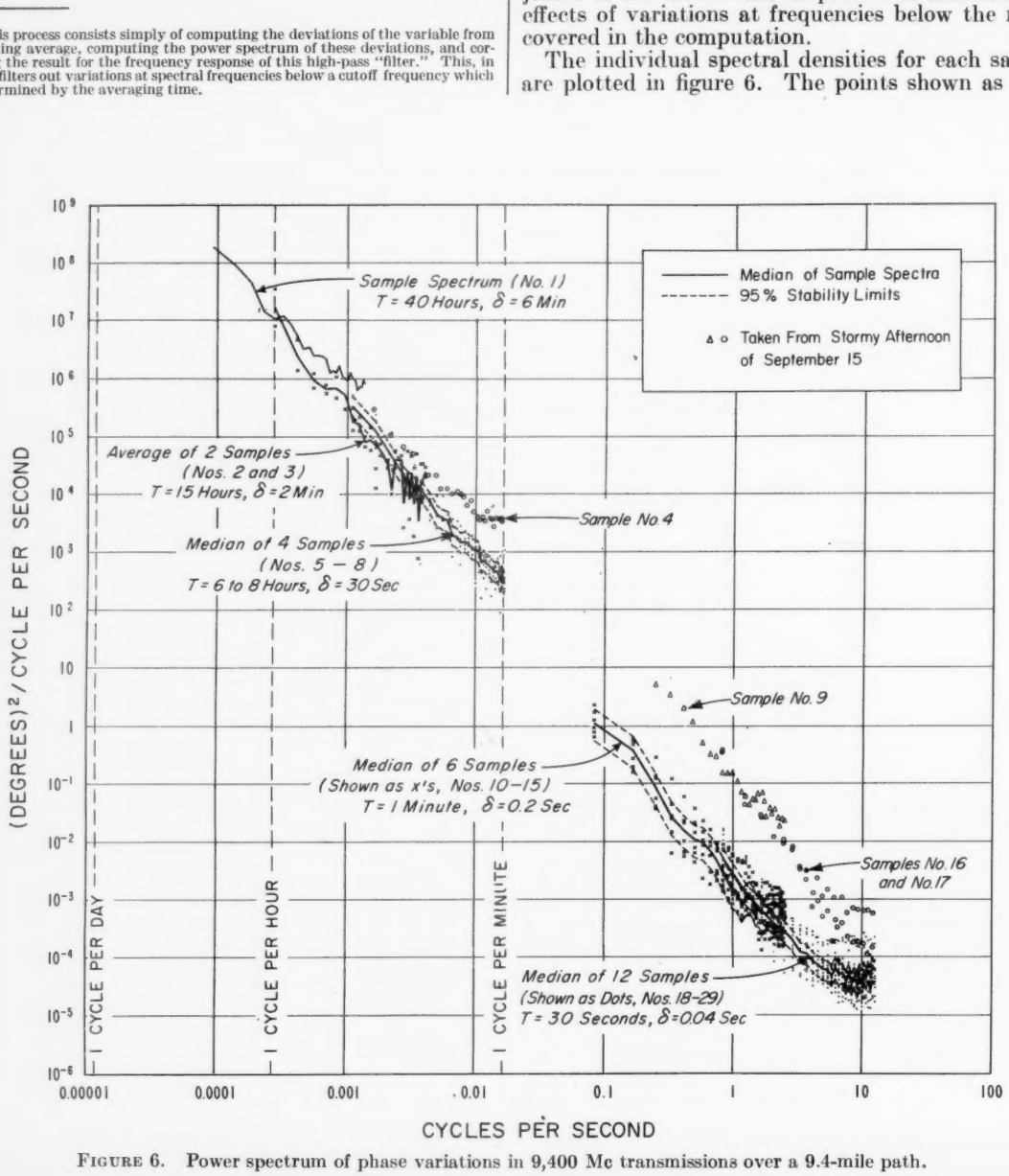


FIGURE 6. Power spectrum of phase variations in 9,400 Mc transmissions over a 9.4-mile path.

dots were computed from data recorded during the stormy afternoon of September 15. The phase recording during that period was characterized by rapid fluctuations which did not recur during the remainder of the experiment. The corresponding spectra were sufficiently different from all the subsequent ones to warrant distinguishing them from the rest. The solid lines join the medians of the sample spectral densities for each group of spectra, excluding those shown as open dots. The range of the points at any given frequency is determined both by sampling fluctuations and by more or less long-term changes in propagation conditions. Some measure of the relative magnitude of these two effects can be obtained by considering the distribution of spectral densities to be expected if the samples had been taken from the same statistical population, i.e., if the time variations in phase constituted a stationary process. In that case the spectral density estimates at each frequency would be distributed according to the chi-squared distribution with an effective number of degrees of freedom determined by the number of data points used in each estimate; the number of frequencies for which estimates were computed; and, to a very limited extent, by the slope of the spectrum. Using the sample median to approximate the "true" median, the dashed lines labelled "95 percent stability limits" represent the bounds within which 95 percent of all such spectral estimates might be expected to fall if the process were indeed stationary. In view of the large long-term changes in phase and meteorological conditions shown in figures 3 and 4, it is not surprising that about 50 percent of the spectral estimates fall outside the 95 percent stability limits, indicating systematic changes in spectral density over the 40-hr period. In fact, it is remarkable that the sample spectra are so similar, both in slope and magnitude.

It will be noted that there is a tendency for the slope to approach zero at the high frequency end of each set of spectra. This is caused by "aliasing", or the addition to the computed densities of power actually present at the higher frequencies For spectra having steep slopes (such as these) aliasing causes an error which is about two to one at the highest frequency of each spectrum and which diminishes rapidly with frequency. Additional flattening of the spectrum above about 6 cps is probably caused by "noise" introduced in the process of reading and analyzing the data, and should be ignored. Making allowances for aliasing and reading error noise, it appears that the slope of the phase spectrum is essentially constant over the range from 8 cycles per day to 12.5 cps, the power density being approximately proportional to $f^{-2.8}$.

Figures 7 and 8 show the seven spectra from the lower frequency range of the Sanborn data and the five spectra from the high-frequency range of the Varian data identified individually by time of day. In general, there appears to be no outstanding correlation between these spectra and either the time of day or gross meteorological variations (with the obvious exception of the spectra taken from the

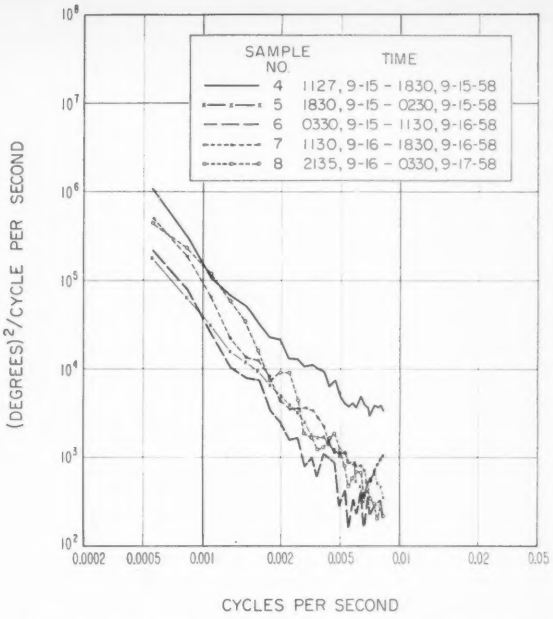


FIGURE 7. *Sample power spectra of phase variations in 9,400 Mc transmissions over a 9.4-mile path.*

Low-pass Varian recording.

stormy afternoon of September 15). One is tempted to attach some significance to the fact that the phase variance was somewhat lower during the period of high humidity following the storm than it was later during the period of warmer and drier weather. This is, of course, too small a data sample to warrant drawing any conclusions, and the correlation of phase and meteorological data will be more thoroughly studied in subsequent measurements.

During the afternoon of September 17 (after the 40-hr run), the polarization of the antennas at both ends of the path was changed from horizontal to vertical and back again to horizontal. A spectrum analysis of samples of phase variations before and after the changes indicated that polarization has no apparent effect on the phase spectrum. This lends some support to the belief that there is little possibility of multipath effects caused by ground reflections from the mesa at Roberts Lake (see figs. 1 and 2).

In summary, the phase spectrum analysis described here indicates that: (1) The slope of the power spectrum is nearly constant over the range of frequencies from 0.0001 cps to about 10 cps, the spectral densities being roughly proportional to $f^{-2.8}$; (2) except for the afternoon of September 15 when a rainstorm crossed the path and caused large high frequency fluctuations in phase, the power spectrum remained fairly constant as a function of time, both with regard to magnitude and to slope; and (3) the phase variations appear to be insensitive to changes in polarization.

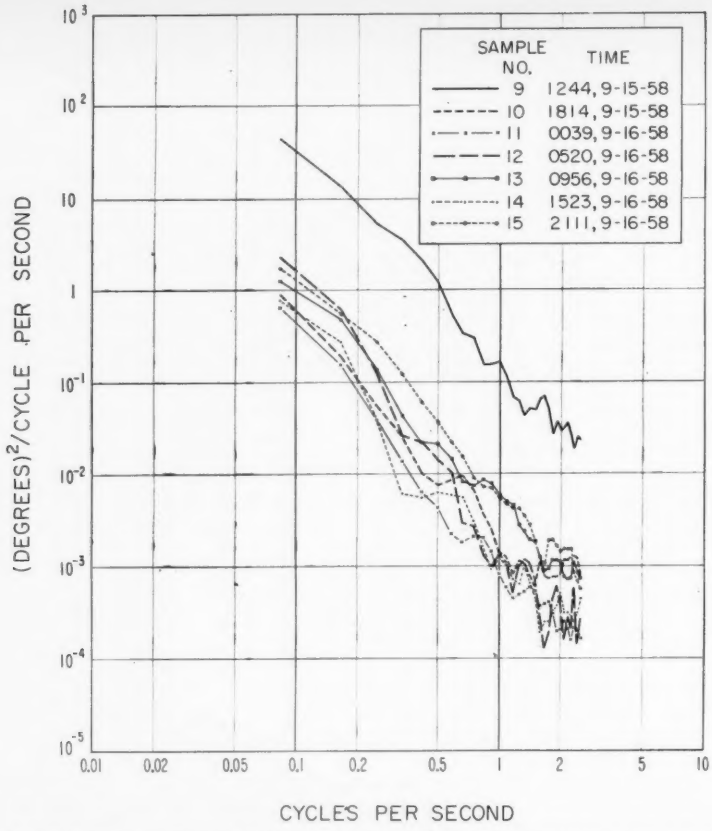


FIGURE 8. *Sample power spectra of phase variations in 9,400 Mc transmissions over a 9.4-mile path.*

Band-pass Sanborn recording

5. Power Spectrum of Radiofrequency Deviations

The time variations in phase can be interpreted as deviations in the radiofrequency of the received signal, and a simple relationship exists between the power spectra of these two parameters. To express the effects of the atmospheric turbulence on the propagated signal in terms of its frequency-modulation effects rather than phase variations, we may write the transmitted and received waves (ignoring attenuation) as follows:

$$e_T = \epsilon^{i\omega t} \tag{1}$$

$$e_R = \epsilon^{i(\omega t - kl)} = \epsilon^{i(\omega t - \theta)}. \tag{2}$$

If we define the frequency of the received signal as the time derivative of the total phase term we obtain:

$$\omega' = \omega - \frac{\partial \theta}{\partial t} \tag{3}$$

or

$$\Delta\omega = \omega - \omega' = \frac{\partial \theta}{\partial t} \tag{4}$$

Thus, the frequency deviation due to the turbulence is proportional to the slope of the phase records obtained in the experiment.

If θ is some function of time which can be represented by the Fourier series:

$$\theta(t) = a_0 + a_1 \cos 2\pi f_1 t + b_1 \sin 2\pi f_1 t + \ldots + a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t + \ldots, \tag{5}$$

its power spectral density at the spectral frequency f_n is proportional to $(a_n^2 + b_n^2)$. Similarly, the frequency deviations can be represented by the series:

$$\frac{\partial \theta}{\partial t} = -2\pi f_1 a_1 \sin 2\pi f_1 t + 2\pi f_1 b_1 \cos 2\pi f_1 t + \ldots - 2\pi f_n a_n \sin 2\pi f_n t + 2\pi f_n b_n \cos 2\pi f_n t + \ldots,$$

and its spectral density at f_n is proportional to $(2\pi f_n)^2(a_n^2 + b_n^2)$.

Figure 9 shows the frequency variation spectra corresponding to the phase spectra shown in figure 6. They were obtained simply by multiplying the

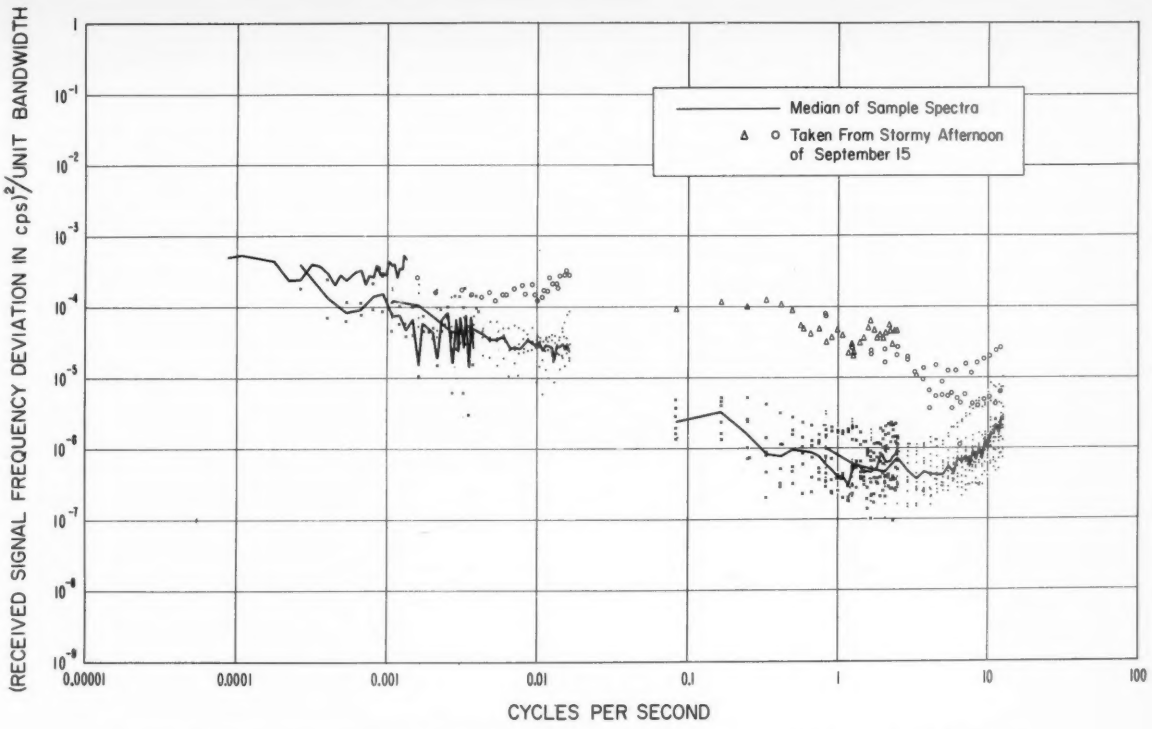


FIGURE 9. *Power spectrum of radiofrequency deviations in 9,400 Mc transmissions over a 9.4-mile path.*

phase spectral density at the frequency f cycles per second by $\left(\frac{\pi f}{180}\right)^2$. The spectral density of radiofrequency variations decreases relatively slowly with increasing spectral frequency, being approximately proportional to $f^{-0.8}$. (The exponent here is simply the value obtained for the phase spectrum increased by two.) The apparent increase in power at the high frequency end (above about 6 cps) is caused by the reading error noise mentioned previously and should be ignored.

6. Comparison of Phase and Atmospheric Refractivity Variations

Variations in the phase of the received signal represent variations in the electrical length of the propagation path, which is proportional to the atmospheric refractivity integrated along the path. Hence, the long-term changes in refractivity measured at a point on the path should be correlated with the long-term changes in electrical path length as deduced from the phase records. This correlation is illustrated in figure 4, in which the refractivity values computed from temperature, pressure, and humidity recordings at each terminal are plotted with the long-term phase variations read from the Varian record. It is perhaps more clearly shown in figure 5, in which time variations in the average of the refractive index at the path terminals and the changes in electrical path length are both expressed in parts per million. This illustrates the fact that the two variables are not only closely correlated (correlation coefficient: 0.915), but their fluctuations agree in magnitude as well.

The following personnel participated in the experiment reported here: J. T. Collins, F. E. Freethey, W. B. Grant, D. P. Harris, H. B. Janes, A. W. Kirkpatrick, D. Smith, M. J. Vetter, and D. M. Waters. Project leader was Moody C. Thompson, Jr.

The assistance of J. W. Herbstreit and K. O. Hornberg is also gratefully acknowledged.

BOULDER, COLO. (Paper 63D1–7).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

System Loss in Radio Wave Propagation

Kenneth A. Norton

(February 26, 1959)

A summary is presented of the ways in which the concept of system loss and the closely related concepts of transmission loss, basic transmission loss, propagation loss, and path antenna gain may be used for precise, yet simple, descriptions of some of the characteristics of radio wave propagation which are important in the design of radio systems. Definitions of various terms associated with the concept of system loss are given which introduce a greater flexibility into its use without any loss in precision. It is shown that the use of these added terms and concepts makes feasible the extension of the use of this method of description to any portion of the radio spectrum. A more general formula for the system loss is given which may be used for antennas with an arbitrarily small separation. Using this formula it is shown that the system loss between small electric or magnetic dipoles separated by a distance $d \ll \lambda$ can be made arbitrarily small even though the individual antennas have large circuit losses. Formulas are developed for the percentage of time that a desired signal is free of interference, and these are used to demonstrate methods for the efficient use of the spectrum. In particular, contrary to general belief, it is shown that efficiency is promoted by the use of high power and high antennas and, in the case of a broadcast service, sufficiently *small* separations so that there is appreciable mutual interference. An analysis is made of the variance of the path antenna gain in ionospheric scatter propagation. Methods are given for the calculation of the transmission loss for the ground wave and tropospheric scatter modes of propagation through a turbulent model atmosphere with an exponential gradient. Examples of such calculations are given which cover a wide range of frequencies and antenna heights. Finally, examples are given of the expected range of various tropospheric point-to-point scatter systems such as an FM multichannel teletype system, a television relay or an FM broadcast relay.

1. Introduction

Although the idea of a radio circuit transmission loss had been in use by engineers concerned with the design of communication systems for many years prior to that time, it was in a paper by the author [1] in 1953 that its great advantages, particularly in connection with ionospheric or tropospheric scatter systems, were first explicitly pointed out. Since that time this concept has been used extensively in the radio propagation studies at the Central Radio Propagation Laboratory and elsewhere, and a fairly large body of conventions has grown up around this usage. It is the purpose of this paper to describe the precise meaning attached to these conventions by the engineers at CRPL, with the hope that these usages will prove equally useful and will be adopted in other laboratories throughout the world. In addition, its use is illustrated by numerous examples throughout the usable portion of the radio spectrum ranging from 3 to 100,000,000 kc.

2. Definitions of System Loss, Transmission Loss, Basic Transmission Loss, Path Antenna Gain, and Path Antenna Power Gain

The *system loss* of a radio circuit consisting of a transmitting antenna, receiving antenna, and the intervening propagation medium is defined as the dimensionless ratio, p_t/p_a, where p_t is the radio-frequency power input to the terminals of the transmitting antenna and p_a is the resultant radio-frequency signal power available at the terminals of the receiving antenna. Both p_t and p_a are expressed in watts. The system loss is usually expressed in decibels:

$$L_s = 10 \log_{10} (p_t/p_a) = P_t - P_a \qquad (1)^1$$

Note that the system loss, as above defined, excludes any transmitting or receiving antenna transmission line losses since it is considered that such losses are readily measurable. On the other hand, the system loss includes all of the losses in the transmitting and receiving antenna circuits, including not only the transmission loss due to radiation from the transmitting antenna and reradiation from the receiving antenna, but also any ground losses, dielectric losses, antenna loading coil losses, terminating resistor losses in rhombic antennas, etc. The inclusion of all of the antenna circuit losses in the definition of system loss provides a quantity which can always be accurately measured and which is directly applicable to the solution of radio system design problems.

For many applications, however, particularly in the study of radio wave propagation, it is convenient to have a definition of a system loss which excludes all of the antenna circuit losses except those associated with the antenna radiation resistances; thus, the *transmission loss* of a radio circuit consisting of a transmitting antenna, receiving antenna, and the intervening propagation medium is defined as the dimensionless ratio, p_t'/p_a', where p_t' is the radio-frequency power radiated from the transmitting antenna, and p_a' is the resultant radiofrequency signal power which would be available from the

[1] Throughout this paper, capital letters are used to denote the ratios, expressed in decibels, of the corresponding quantities designated with lower-case type; e.g., $P_t = 10 \log_{10} p_t$ is the input power to the transmitting antenna expressed in decibels above 1 w.

receiving antenna if there were no circuit losses other than those associated with its radiation resistance. The transmission loss is usually expressed in decibels:

$$L \equiv 10 \log_{10}(p_t'/p_a') = L_s - L_{tc} - L_{rc}, \quad (2)$$

where L_{tc} and L_{rc} are the losses, expressed in decibels, in the transmitting and receiving antenna circuits, respectively, excluding the losses associated with the antenna radiation resistances; i.e.,

$$L_{tc} = 10 \log_{10}(r_t'/r_t) = 10 \log_{10}(1+\Delta_{tc})$$

and

$$L_{rc} = 10 \log_{10}(r_r'/r_r) = 10 \log_{10}(1+\Delta_{rc}),$$

where r' is the actual resistive component of the antenna circuit, r is the radiation resistance, and the subscripts t and r refer to the transmitting and receiving antennas, respectively.

In order to separate the effects of the transmitting and receiving antenna gains and circuit losses from the effects of the propagation, it is convenient to define the *basic transmission loss*, L_b (sometimes called path loss) as the transmission loss expected between fictitious loss-free isotropic transmitting and receiving antennas at the same locations as the actual transmitting and receiving antennas. This serves also to define the *path antenna directive gain*, G_p:

$$G_p \equiv L_b - L, \quad (3)$$

$$L_b = L + G_p = L_s + G_p - L_{tc} - L_{rc}. \quad (4)$$

In some cases it may be quite difficult to measure the antenna circuit losses; thus it is convenient to define the *path antenna power gain*, G_{pp}, as

$$G_{pp} \equiv L_b - L_s = G_p - L_{tc} - L_{rc}. \quad (5)$$

It is seen that the path antenna power gain is the change in the system loss when loss-less isotropic antennas are used at the same locations as the actual antennas; note that G_{pp} will be negative when the antenna circuit losses exceed the path antenna directive gain. Throughout the remainder of this paper the term path antenna gain and symbol G_p are often used when a distinction between the directive gain and the power gain is unnecessary.

In some idealized situations the path antenna power gain, G_{pp}, is simply the sum $(G_{tp}+G_{rp})$ of the free space power gains G_{tp} and G_{rp} of the transmitting and receiving antennas relative to loss-less isotropic antennas. However, in most practical situations G_{pp} is less than $G_{tp}+G_{rp}$ because of the complex nature of the received field. The path antenna power gain may be measured by determining the increase in the system loss when both the transmitting and receiving antennas are replaced *simultaneously* by simple standard antennas such as short electric or magnetic dipoles, and then adding the calculated path antenna power gain corresponding to the use of the standard antennas. In the case of ionospheric or tropospheric scatter propagation, the path antenna power gain is sometimes substantially smaller than the sum of the free space power gains G_{tp} and G_{rp}. In such cases the path antenna power gain cannot be determined accurately as the sum of the effective power gains of the transmitting and receiving antennas (as determined by replacing first one antenna by a standard antenna and then the other antenna by a standard antenna) since such effective power gains depend upon the gain of the antenna used at the other terminal.

In the case of ionospheric or tropospheric propagation, the transmission loss L, the basic transmission loss L_b, and the path antenna gain G_p, are all random variables with respect to time and tend to be normally distributed about their expected values. Furthermore, L and G_p are typically negatively correlated with each other, and thus the variance of L_b is usually substantially less than the sum of the variances of L and G_p (see sec. 9); for this reason it will often be more practical simply to measure the transmission loss with the particular antennas intended for use rather than attempt to calculate the expected transmission loss and its variance with time in terms of the measured or calculated values of the basic transmission loss and the path antenna gain.

Note also that the path antenna gain may actually be negative. For example, the path antenna gain will usually be negative for ground wave or tropospheric wave propagation between a vertically polarized and a horizontally polarized antenna, and the concept of path antenna gain should prove to be useful for expressing the results of such cross polarization measurements.

3. Transmission Loss in Free Space

As an example of the simplicity of transmission loss calculations in some cases, we may consider the transmission loss between two isotropic antennas in free space. At a distance, d, very much greater than the wavelength, λ, the field intensity, expressed in watts per square meter, is simply $p_t'/4\pi d^2$ since the power is radiated uniformly in all directions. Since the effective absorbing area of the receiving antenna is $\lambda^2/4\pi$, the available power at the terminals of the loss-free isotropic receiving antenna is given by

$$p_a' = (\lambda^2/4\pi)\cdot(p_t'/4\pi d^2). \quad (6)$$

Consequently, the basic transmission loss in free space may be expressed

$$L_{bf} = 10 \log_{10}(4\pi d/\lambda)^2 \qquad (\text{free space}; d >> \lambda). \quad (7)$$

Since the free space gain of a short electric loss-less dipole is $g_t = g_r = 1.5$, the path antenna gain for two optimally oriented short electric loss-less dipoles in free space is

$$G_p = G_t + G_r = 3.52 \text{ db}. \quad (8)$$

Consequently, the transmission loss between two optimally oriented short electric loss-less dipoles in free space is

$$L = 10 \log_{10}(4\pi d/\lambda)^2 - 3.52. \quad (9)$$

4. Influence of the Antenna Environment and Definition of Propagation Loss

In order to illustrate the influence of changes in the impedances of the antennas caused by environmental factors which may, in part, be independent of the antenna circuit losses, we will consider the transmission loss between two short vertical electric loss-less dipoles at heights h_t and h_r, respectively, above a plane perfectly conducting surface and separated by a distance, d, along the surface large with respect to the wavelength. Although we will treat here in detail only the case of antennas over a perfectly conducting plane, the concepts are identical in the case of any other environmental conditions and only the magnitudes of the effects will be different. In free space the field strength e, expressed in volts per meter at a distance d, expressed in meters, in the equatorial plane of a short electric loss-less dipole radiating p_t' watts is given by

$$\frac{e^2}{\eta_0}=\frac{1.5p_t'}{4\pi d^2}, \tag{10}$$

where $\eta_0=4\pi c\cdot 10^{-7}=$impedance of free space expressed in ohms, and $c=2.997925\cdot 10^8$ m/sec= velocity of light in free space. Note that the factor (1.5) is just the free space gain of the transmitting dipole antenna. The radiation resistance of a short vertical electric dipole of effective length l and at a height h_t above a perfectly conducting plane is given by

$$r=\frac{2\pi\eta_0 l^2}{3\lambda^2}\,[1+\Delta_t]=r_f[1+\Delta_t], \tag{11}$$

$$\Delta_t=\frac{3}{(2kh_t)^2}\left[\frac{\sin(2kh_t)}{2k\,h_t}-\cos(2k\,h_t)\right]. \tag{12}$$

In the above $k=2\pi/\lambda=2\pi f/c$, i.e., λ is the wavelength in free space. Note that Δ_t approaches zero at large heights above the surface and r approaches its free space value, r_f. On the other hand, $\Delta_t=1$ for $h_t=0$, and the radiation resistance is then just twice its free space value. The field intensity expressed in watts per square meter for a short vertical electric loss-less dipole over the perfectly conducting surface may be expressed

$$\frac{e^2}{\eta_0}\cong\frac{1.5p_t'[2\cos^2\psi\cos(k\,h_t\sin\psi)]^2}{4\pi d^2[1+\Delta_t]}. \tag{13}$$

In the above expression $\tan\psi=h_r/d$ and the distance d along the surface must be large with respect to both λ and h_t; in this case the distance between the antennas is approximately $d/\cos\psi$. Equations (11), (12), and (13) were derived by Schelkunoff in chapters VI and IX of [2]. Since $\Delta_t=1$ for $h_t=0$, the field intensity is 3 db greater when $\psi=0$ and the dipole is on the surface of a perfectly conducting plane (i.e., $e^2/\eta_0=3p_t'/4\pi d^2$) than when it is in free space; note that in free space (10) must be used rather than (13), since the ground reflection influences the radiation for all values of h_t. In more familiar units, when $h_t=h_r=0$, (13) may be expressed

$$e(\mu v/m)=2.998962\cdot 10^5\sqrt{p_t'(kw)}/d(km). \tag{14}$$

The effective absorbing area of a short vertical electric loss-less dipole receiving antenna at a height h_r above the perfectly conducting plane may be expressed

$$a_e=\frac{1.5\lambda^2\cos^2\psi}{4\pi[1+\Delta_r]}, \tag{15}$$

where Δ_r is defined by (12) with h_t replaced by h_r. Since $p_t'=e^2a_e/\eta_0$ we find by combining (13) and (15) that the transmission loss between two short vertical electric loss-less dipoles at heights h_t and h_r above a plane perfectly conducting surface may be expressed

$$L\cong 10\log_{10}\left\{\frac{[4\pi d/\lambda\cos\psi]^2[1+\Delta_t][1+\Delta_r]}{(1.5)^2[2\cos^2\psi\cos(kh_t\sin\psi)]^2}\right\} \tag{16}$$

or

$$L_s=L=L_{bf}-G_p-6.02+A+L_t+L_r\equiv L_p+L_t+L_r, \tag{17}$$

where $L_{t,r}=10\log_{10}[1+\Delta_{t,r}]$, $G_p\cong 20\log_{10}[(3/2)\cos^2\psi]$ and $A\cong -20\log_{10}[\cos(kh_t\sin\psi)]$. Note that $L_s=L$ in this case since there are no losses other than radiation losses. The factor L_t is shown graphically as a function of (h/λ) in figure 1. It is of interest to note that the transmission loss between the dipoles on the plane perfectly conducting surface, $h_t=h_r=0$, is the same as if the dipoles were separated by the same distance in free space, although the field intensity is 3 db greater for the same power radiated. On the other hand, when the dipoles are several wavelengths above the perfectly reflecting surface ($h_t=h_r>>\lambda$) and are separated by a large distance ($\psi\cong 0$), the transmission loss is 6.02 db less than for dipoles separated by the same distance in free space. Equation (17) illustrates the definition of L_p, the *propagation loss*, i.e., the transmission loss expected if the antennas had gains and circuit resistances the same as if they were in free space. More generally, L_t and L_r are defined as

$$L_{t,r}=10\log_{10}(r'/r_f), \tag{18}$$

where r' is the actual resistance of the antenna in the presence of its environment and r_f is the radiation resistance it would have had if it were in free space. When we note that L_t and L_r may vary substantially from one antenna installation to the next, depending upon the circuit losses, polarization, ground conditions, whether or not a ground screen is used, and upon other environmental factors such as the presence of trees or overhead wires, it becomes clear why it is desirable to separate these components from the system loss and to have a

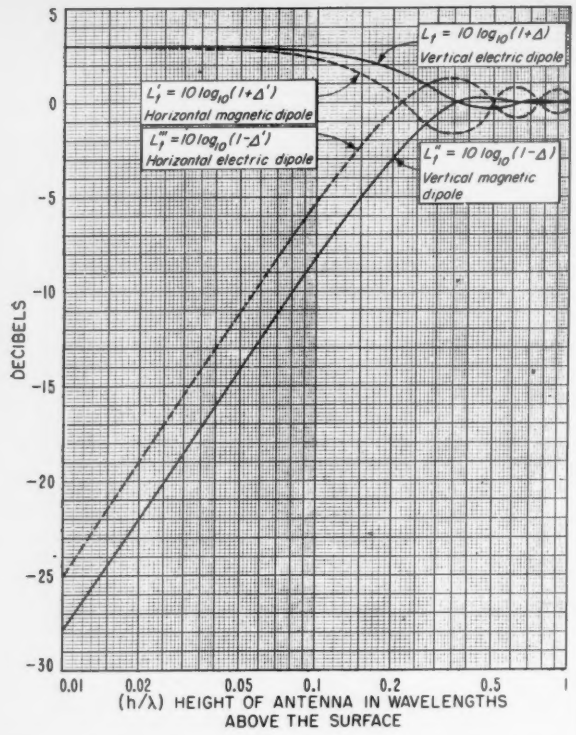


FIGURE 1. *Transmission loss arising from a change in the radiation resistance of short dipole antennas near a perfectly-conducting surface.*

propagation loss, L_p, independent of these antenna environmental conditions.

$$L_p \equiv L_s - L_t - L_r \tag{19}$$

The above is not the only logically consistent method of describing the gains and losses of antennas in the presence of an environment. Thus Schelkunoff [2] considered the perfectly reflecting surface to be an integral part of the transmitting and receiving antennas, and set $G_{ts}=10 \log_{10}\{3/[1+\Delta_t]\}$ and $G_{rs}=10 \log_{10}\{3/[1+\Delta_r]\}$ where the subscript s refers to Schelkunoff's usage. In other words, he referred his maximum gains to those expected for isotropic antennas with the earth removed; this leads to a path antenna gain, referred to that expected between isotropic antennas with the earth removed, given by $G_{ps}=G_p+6.02-L_t-L_r-A$, and his tranmission loss is then simply $L=L_{bf}-G_{ps}$. However, this method of approach is *not recommended* since (a) it is impracticable to remove the earth in order to measure G_{ps} and (b) it would lead to antenna gains 3 db larger than their free space values even when they are many wavelengths above a perfectly reflecting plane surface, and this is inconsistent with the present usage of the concept of antenna gain in the higher frequency ranges.

Although it appears to be desirable to separate the effects of earth reflections and sometimes other environmental effects from the gains of the antennas by means of the propagation loss concept described earlier, there will be other situations in which such a separation of environmental effects is undesirable. For example, the power gain of an antenna mounted on an aircraft, satellite, or space vehicle should be considered as the gain which would be determined by comparison of the fields obtained in various directions as the vehicle is rotated in free space with those obtained from a standard antenna with the vehicle removed; i.e., in these cases the vehicle is to be considered an integral part of the antenna.

Suppose now that we use small loop antennas of area S, with their axes normal to the plane of propagation, parallel to the perfectly conducting surface and at heights h_t and h_r, respectively. In this case the electric vector lies in the horizontal plane and the radiation resistance is [2]

$$r_m' = \frac{8\pi^3 \eta_0 S^2}{3\lambda^4}[1+\Delta_r'] = r_{mf}[1+\Delta_r']. \tag{20}$$

The subscript m refers to the magnetic dipole, and the prime to the effect of the surface.

$$\Delta_r' = (3/2)\left[\left(1-\frac{1}{(2kh_r)^2}\right)\frac{\sin(2kh_r)}{2kh_r}+\frac{\cos(2kh_r)}{(2kh_r)^2}\right], \tag{21}$$

$$a_e = \frac{\lambda^2(3/2)}{4\pi[1+\Delta_r']}, \tag{22}$$

$$G_p' = 20 \log_{10}[3/2] = 3.52, \tag{23}$$

$$A' = -20 \log_{10}[\cos(kh_t \sin\psi)], \tag{24}$$

$$L_{t,r}' = 10 \log_{10}[1+\Delta_{t,r}']. \tag{25}$$

The factor $L_{t,r}'$ is also shown graphically on figure 1. Note that Δ_r' approaches zero at large heights and $\Delta_r'=1$ for $h_r=0$.

Consider next the tranmission loss between two small loop antennas at heights h_t and h_r, respectively, above a perfectly conducting surface with their axes normal to this surface. In this case the electric vector lies in the horizontal plane and the radiation resistance is [2]

$$r_m'' = \frac{8\pi^3 \eta_0 S^2}{3\lambda^4}[1-\Delta] = r_{mf}[1-\Delta], \tag{26}$$

$$G_p'' = 20 \log_{10}[(3/2)\cos^2\psi], \tag{27}$$

$$A'' = -20 \log_{10}[\sin(kh_a \sin\psi)], \tag{28}$$

$$L_{t,r}'' = 10 \log_{10}[1-\Delta_{t,r}]. \tag{29}$$

The factor $L_{t,r}''$ is also shown as a function of (h/λ) on figure 1.

Finally, consider the transmission loss between two short horizontal electric dipoles of effective

length l, normal to the plane of propagation and at heights h_t and h_r, respectively, above a perfectly conducting plane surface. In this case [2],

$$r''' = \frac{2\pi\eta_0 l^2}{3\lambda^2}[1-\Delta'] = r_f[1-\Delta'], \tag{30}$$

$$G'''_p = 20\log_{10}(3/2) = 3.52, \tag{31}$$

$$A''' = -20\log_{10}[\sin(kh_a \sin\psi)], \tag{32}$$

$$L'''_{t,r} = 10\log_{10}[1-\Delta'_{t,r}]. \tag{33}$$

The factor $L'''_{t,r}$ is also shown graphically on figure 1. Note that L''' and L'' both approach $(-\infty)$ as h approaches zero, but the radiation resistance simultaneously approaches zero, and it would be difficult in practice to keep the radiated power constant as the antennas are brought nearer and nearer to the surface. When h_t and h_r are both much less than a wavelength, $A''+L''_t+L''_r$ for horizontal loops becomes independent of these heights and equal to $20\log_{10}(kd/10)$; similarly $A'''+L'''_t+L'''_r$ for horizontal electric dipoles approaches $20\log_{10}(kd/5)$ for h_t and h_r much less than a wavelength.

All of the above specific results refer to the case of a perfectly conducting plane surface and to distances $d \gg \lambda$. For a finitely conducting ground, the attenuation A in the above expressions will be substantially modified. For example, for vertical electric dipoles over a flat earth of finite conductivity and with $h_t = h_r = 0$,

$$A = -20\log_{10}|1 - i\sqrt{\pi p}\exp(-p)\operatorname{erfc}(i\sqrt{p})| \tag{34}$$

Here p denotes Sommerfeld's numerical distance as defined for a time factor $\exp(+i\omega t)$. In reference [3] a comprehensive discussion is given of the radiation fields of electric and magnetic dipoles over a finitely conducting plane earth; in this reference a time factor $\exp(-i\omega t)$ was used, but in the present paper $\exp(+i\omega t)$ is used in order to conform with current engineering practice. Furthermore, Δ and Δ' will be modified when the antennas are located over a finite ground [4, 5, 6, 7], but this difference will largely be cancelled if a large ground screen is used under the antennas. Although (26) indicates that r''_m approaches zero as the vertical magnetic dipoles approach the perfectly conducting surface, Wait [4] has shown that r''_m actually becomes very large when such loops are brought near a finitely conducting ground.

5. Relation Between Propagation Loss, L_p, and Field Strength, E

The results of many previous studies of radio wave propagation in the low and medium frequency range have been expressed in terms of the field strength E, expressed in decibels above 1 μv/m for an unattenuated field I, at a unit distance, i.e., to the radiation field expected at a unit distance on a perfectly conducting plane surface. It is convenient to express I in decibels above 1 v, e.g., 1 v/m at 1 m. We will relate such values to the propagation loss, L_p.

In free space the inverse distance field I_f may be obtained from (10); when we note that $P_t' = P_t - L_t$,

$$I_f = 10\log_{10}(p_t' c\cdot 10^{-7} g_t) = P_t - L_t + G_t + 14.77. \tag{35}$$

Over a perfectly conducting ground plane the inverse distance field I, may be obtained from (13):

$$I = P_t - L_t + G_t + 20.79. \tag{36}$$

The relation between the available power, p_a, from the receiving antenna and the field strength, e, may be expressed in decibels as follows:

$$E = 10\log_{10}(4\pi\eta_0 p_a[r'/r_f]\cdot 10^{12}/\lambda^2 g_r)$$
$$= P_a + 20\log_{10} f_{\mathrm{Mc}} - G_r + L_r + 107.22. \tag{37}$$

If we solve (36) for P_t and (37) for P_a and combine the results, we obtain the following expression:

$$L_s = P_t - P_a = I - E_I - G_p + 20\log_{10} f_{\mathrm{Mc}} + 86.43 + L_t + L_r \tag{38}$$

In the literature E is sometimes referred to an inverse distance field of 300 v, i.e., 3×10^5 μv/m at 1 km (i.e., by (14) to 1 kw radiated from a short vertical electric dipole on a perfectly conducting plane surface) and in this case $I = 49.54$ db so that

$$L_p = L_s - L_t - L_r = 135.97 - G_p + 20\log_{10} f_{\mathrm{Mc}} - E_{I=49.54} \tag{39}$$

In the higher frequency ranges, however, E is usually referred to an inverse distance field in free space of 222 v (i.e., by (35) to 1 kw radiated from a half wave dipole in free space) and in this case $I_f = 46.92$ db; if we now solve (36) for $P_t - L_t$, solve (37) for $P_a + L_r$, we obtain the following relation:

$$L_p = P_t - L_t - P_a - L_r = 139.37 - G_p + 20\log_{10} f_{\mathrm{Mc}} - E_{I_f=46.92} \tag{39a}$$

One of the important advantages of expressing propagation results in terms of the propagation loss L_p rather than in terms of E is the fact that L_p is dimensionless and so does not require a specification of the reference inverse distance field or of the radiated power.

6. System Loss, Transmission Loss, Propagation Loss, and Relative Phase of Currents in Two Antennas Expressed in Terms of Their Self and Mutual Impedances

All of the preceding analyses are limited to the case where the antennas are sufficiently far apart ($d \gg \lambda$) so that the magnitude of their mutual im-

pedance is small compared to the self resistances of each of the antennas. In this section a completely general expression will be obtained for the system loss. It is convenient to use the T-network of figure 2 to represent the impedances of the antennas. The accessible terminals of the transmitting antenna a and of the receiving antenna b are denoted by AA and BB, respectively. The central member of the T is the mutual impedance z_m between the two antennas while z_a and z_b are the self-impedances of the two antennas; thus the impedance at the terminals AA, with BB open circuited, is z_a while the impedance at the terminals BB, with AA open circuited, is z_b.

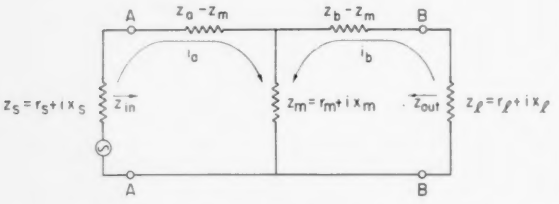


FIGURE 2. *T-network for system loss analysis.*

Let z_l indicate the receiving antenna load impedance at BB. From figure 2 it is easy to show that the impedance z_{in} at AA is given by

$$z_{in}=z_a-\frac{z_m^2}{z_b+z_l}. \tag{40}$$

Since there are no source voltages in the mesh in which the current i_b is flowing, we may write

$$i_b(z_b+z_l)+i_a z_m=0. \tag{41}$$

The above relation provides a means for determining the relative magnitude and phase of the current i_b in antenna b relative to the magnitude and phase of the current i_a in antenna a. The fact that i_a and i_b have the same direction of flow through z_m may be established by considering the limiting case as the spacing between the two antennas approaches zero. The ratio of the power $|i_a^2|$ Re z_{in} from the source to the power $|i_b^2|$ Re z_l delivered to the load may now be expressed

$$\frac{|i_a^2|\ \mathrm{Re}\ z_{in}}{|i_b^2|\ \mathrm{Re}\ z_l}=\frac{\mathrm{Re}\ z_{in}}{\mathrm{Re}\ z_l}\left|\frac{z_b+z_l}{z_m}\right|^2. \tag{42}$$

Note that the maximum power will be delivered to the load, i.e., the above ratio will be minimized for a constant power input to the transmitting antenna, when the load impedance z_l is adjusted so that it is equal to the complex conjugate z_{out}^* of the output impedance $z_{out}=z_b-z_m^2/(z_a+z_s)$ where z_s is the impedance of the source. In this case i_b^2 Re z_l is equal to the available power from the receiving antenna and the system loss is equal to

$$L_s=10\log_{10}\left[\frac{r_a'|z_b+z_{out}^*|^2}{\mathrm{Re}\ z_{out}|z_m^2|}-\frac{|z_b+z_{out}^*|^2}{|z_m^2|}-\frac{|z_b+z_{out}^*|\cos(2\phi-\theta)}{\mathrm{Re}\ z_{out}}\right], \tag{43}$$

where r_a' and r_b' denote the real parts of the antenna self impedances z_a and z_b, $z_m\equiv|z_m|\exp(i\phi)$ and $(z_b+z_{out}^*)=2r_b'-[z_m^2/(z_a+z_s)]^*\equiv|z_b+z_{out}^*|\exp(i\theta)$. Equation (43) is the most general expression for the system loss; this same equation may be used for calculating the transmission loss L simply by replacing the actual antenna resistances r_a' and r_b' by the radiation resistances r_a and r_b; furthermore, (43) gives the propagation loss L_p if we replace r_a' and r_b' by the radiation resistances r_{fa} and r_{fb} which the antennas would have if they were located in free space.

Note now that $z_{out}\cong z_b$ provided $|z_m^2|\ll|z_b|\cdot|z_a+z_s|$ and this will be the case (a) at large distances or (b) even at short distances if the source impedance z_s is sufficiently large; in this case $z_l=z_{out}^*\cong z_b^*$. Note that this same relation $z_l=z_b^*$ would apply in case the load impedance were matched to the receiving antenna impedance with the terminals of the transmitting antenna open or with the transmitting antenna at a sufficiently large distance so that it has a negligible effect on z_{out}. In this special case (41) and (43) become

$$i_b(z_l=z_b^*)=-i_a\frac{|z_m|\exp(i\phi)}{2r_b'}, \tag{44}$$

$$L_s(z_l=z_b^*)=10\log_{10}\left\{\frac{4r_a'r_b'}{|z_m^2|}-2\cos(2\phi)\right\}. \tag{45}$$

It will be of interest to illustrate the use of (45) by means of two simple examples. We will consider first the system loss between two short electric dipoles, each perpendicular to the line joining their centers, and separated by a distance d in free space. It may be shown [2] that the mutual impedance z_m between the dipoles of effective lengths l_a and l_b, which are each much shorter than a wavelergth, is given by

$$z_m=\frac{i\eta_0 l_a l_b}{2\lambda d}\left[1-\frac{1}{k^2d^2}-\frac{i}{kd}\right]\exp(-ikd)\equiv|z_m|\exp(i\phi). \tag{46}$$

Here $kd=2\pi\ d/\lambda$. The antenna self resistances are given by

$$r_a'=2\pi\eta_0 l_a^2(1+\Delta_{tc})/3\lambda^2, \tag{47}$$

$$r_b'=2\pi\eta_0 l_b^2(1+\Delta_{rc})/3\lambda^2. \tag{48}$$

The factors $(1+\Delta_{tc})$ and $(1+\Delta_{rc})$ in (47) and (48) allow for the circuit losses.

$$|z_m|^2=\frac{\eta_0^2 l_a^2 l_b^2}{4\lambda^2 d^2}\left\{\left(1-\frac{1}{k^2d^2}\right)^2+\frac{1}{k^2d^2}\right\}, \tag{49}$$

$$\phi=\frac{\pi}{2}-kd-\tan^{-1}\left\{\frac{1}{kd}\Big/\left(1-\frac{1}{k^2d^2}\right)\right\}, \tag{50}$$

$$L_s\,(z_l=z_b^*)=10\log_{10}\left[\frac{(2kd)^2\,(1+\Delta_{tc})\,(1+\Delta_{rc})}{(3/2)^2\left\{\left(1-\frac{1}{k^2d^2}\right)^2+\frac{1}{k^2d^2}\right\}}\right.$$
$$\left.-2\cos\,(2\phi)\right]. \tag{51}$$

When the distance $d\gg\lambda$, the second term in (51) is negligible compared with the first term and (51) approaches (9)$+L_{tc}+L_{rc}$ as expected. At the other extreme when $d\ll\lambda$, the first term in (51) is negligible compared to the second term, ϕ approaches $\pi/2$ and $L_s(z_l=z_b^*)=10\log_{10}2=3.01$ db. Thus, in this case, half of the power from the source is available in the load resistance. The above formulation of the transmission loss problem in terms of the antenna self and mutual impedances is due to Wait [8] and (51) is similar to the solution Wait obtained in appendix II of his paper for two vertical dipoles on a flat perfectly-conducting ground plane.

The transmission loss may be obtained from (51) if we set $\Delta_{tc}=\Delta_{rc}=0$ and we find that $L(z_l=z_b^*)$ between short dipoles is then dependent only on kd. Figure 3 gives the transmission loss versus (d/λ) for this special case, and the system loss for the case $(1+\Delta_{tc})=(1+\Delta_{rc})=2$. By the same method used for deriving (51) it may be shown that it applies also to small magnetic dipoles, i.e., to two small loop antennas with the planes of each loop lying in the same plane.

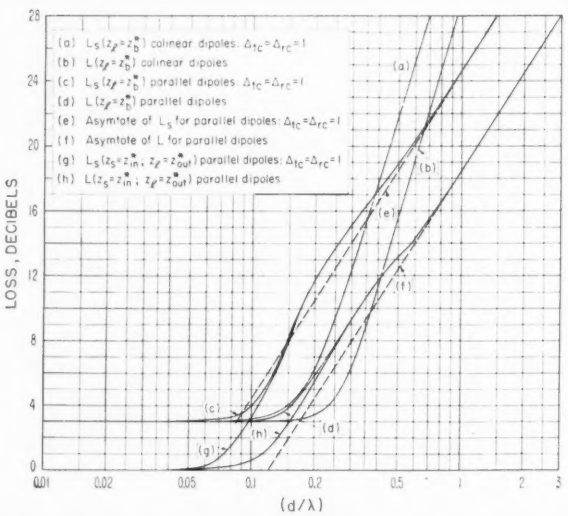


FIGURE 3. *System loss and transmission loss between small electric or magnetic dipoles in free space.*

Consider next two short electric colinear dipoles separated by a distance d in free space. In this case,

$$|z_m|^2=\frac{\eta_o^2 l_a^2 l_b^2}{\lambda^2 d^2 k^2 d^2}\left\{1+\frac{1}{k^2d^2}\right\}, \tag{52}$$

$$\phi=\frac{\pi}{2}-kd+\tan^{-1}(kd), \tag{53}$$

$$L_s(z_l=z_b^*)=10\log_{10}\left[\frac{4k^4d^4\,(1+\Delta_{tc})\,(1+\Delta_{rc})}{9\left\{1+\frac{1}{k^2d^2}\right\}}\right.$$
$$\left.-2\cos\,(2\phi)\right]. \tag{54}$$

Figure 3 also shows the transmission loss, as given by (54) with $\Delta_{tc}=\Delta_{rc}=0$, as well as the system loss for the case $(1+\Delta_{tc})=(1+\Delta_{tr})=2$, between short electric colinear dipoles; (54) also applies to small colinear magnetic dipoles, i.e., small loop antennas with their axes colinear.

Equation (45) and consequently (51) and (54) are not very realistic expressions for the system loss when the distance (d/λ) is small since they are strictly applicable; i.e., the power delivered to the load is the available power, only if z_s increases without limit as (d/λ) decreases. One other example will be used to illustrate the problem. Thus we will assume that the impedance of the source is matched to the transmitting antenna input impedance: $z_s=z_{in}^*$ while the impedance of the load is *simultaneously* matched to the receiving antenna output impedance: $z_l=z_{out}^*$. In this case we may eliminate z_s^* from the two equations: $z_l=z_b^*-(z_m^2)^*/(z_a^*+z_s^*)$ and $z_s^*=z_a-(z_m^2)/(z_l+z_b)$, and in this way find the following expression for $z_l\equiv r_l+iz_l$:

$$z_l=r_b'\left\{1-\frac{|z_m^2|\,\cos\,(2\phi)}{r_a'r_b'}-\frac{|z_m^4|\,\sin^2\,(2\phi)}{4r_a'^2r_b'^2}\right\}^{1/2}$$
$$-ix_b+ir_b'\left\{\frac{|z_m^2|\,\sin\,(2\phi)}{2r_a'r_b'}\right\}. \tag{55}$$

We may determine $z_s\equiv r_s+ix_s$ from (55) simply by interchanging r_a' and r_b' and x_a with x_b. The system loss for this case may now be determined by introducing the above expressions in (42):

$$L_s(z_s=z_{in}^*;z_l=z_{out}^*)=10\log_{10}\left[\frac{2r_a'(r_b'+r_l)}{|z_m^2|}\right.$$
$$\left.-\cos\,(2\phi)\right]. \tag{56}$$

The above is the general expression for the system loss when the source is matched to the input impedance z_{in} and the load is simultaneously matched to the output impedance z_{out}. This is, of course, the most favorable condition for the transfer of power from the source to the load. Equations (55) and (56) are applicable to any kind of antennas and we

see that the system loss depends only on the self resistances r'_a and r'_b of the two antennas and their mutual impedance $z_m \equiv \| z_m \| \exp(i\phi)$.

When $d \gg \lambda$, r_l approaches r'_b and (56) approaches (45) since the second terms in (45) and in (56) are negligible compared to the first terms; thus we conclude that the system loss does not depend upon the condition of matching of the source to the input impedance when d is sufficiently large so that $|z_m^2| \ll r'_a r'_b$.

The behavior of (55) and (56) when $d \ll \lambda$ will depend markedly upon the particular characteristics of the antennas involved. At such short spacings one would normally be interested only in antennas with dimensions small with respect to their spacing and thus small with respect to the wavelength; thus, when $d \ll \lambda$, we are led to consider the special case of short electric or small magnetic dipoles. When (47), (48), (49), and (50) are used in conjunction with (55) and (56) we find that the system loss, $L_s (z_s = z_{in}^*; z_l = z_{out}^*)$, approaches zero when $d \ll \lambda$, i.e., when $(kd)^6 \ll (9/16)/\{(1+\Delta_{tc})(1+\Delta_{rc})\}$, even if r'_a and r'_b are very large with respect to their radiation resistance components; see figure 3 for a graph of the transmission loss and for the system loss with $(1+\Delta_{tc}) = (1+\Delta_{rc}) = 2$ for this type of impedance matching. A system with near zero loss is very remarkable, but may be understood when we note that (r_l/r'_b) and (r_s/r'_a) both increase in this case without limit as d approaches zero; thus the power from the source $i_a^2 r_s$ is much larger than the power $i_a^2 r'_a$ dissipated in the transmitting antenna while simultaneously the power $i_b^2 r_l$ delivered to the load is very much larger than the power $i_b^2 r'_b$ dissipated in the receiving antenna. Although the system loss is nearly equal to zero, only about half (actually a fraction $[r'_b/(r'_a+r'_b)]$) of the power is available in the load while the remaining fraction $[r'_a/(r'_a+r'_b)]$ is dissipated in the impedance of the source.

It might appear that a very efficient short range communication system could be developed based on the above analysis by using sufficiently low frequencies that $\lambda \gg d$; note that the signal-to-noise ratio should be quite favorable since the small antennas used would pick up very little noise power, and this small noise power would be in competition essentially with the large power directly available from the source. Note also that this communication system would create very little interference since the power radiated $i_a^2 r_a + i_b^2 r_b$ is much smaller than the power $i_b^2 r_l^2$ delivered to the load. However, the distance out to which the system loss is less than 3db for such a system is quite small as may be seen from the following equation which applies when $r'_a \gg r_a$ and $r'_b \gg r_b$:

$$d(L_s = 3 \text{ db}) = \frac{\lambda}{2\pi} \left\{ \frac{(9/16) r_a r_b}{r'_a r'_b} \right\}^{1/6}. \qquad (57)$$

If we assume that the antennas are grounded vertical radiators and substitute (11) with $\Delta_t = 1$ for r_a and r_b in (57), we obtain

$$d(L_s = 3 \text{ db}) = 1.68 \lambda^{1/3} l_a^{1/3} l_b^{1/3} / (r'_a r'_b)^{1/6} \qquad (58)$$

If we assume $\lambda = 100{,}000$ m (f = 3 kc), $l_a = l_b = 100$ m, i.e., with actual heights without top loading equal to 200 m, and $r'_a = r'_b = 10$ ohms, then $d(L_s = 3 \text{ db}) = 781$ m; since very good ground systems would be required to reduce r'_a and r'_b to only 10 ohms, it appears that such a system has little promise. This statement should not be interpreted to mean that low frequencies are not useful for short range communication systems, but merely that such systems with less than 3-db system loss will have a very limited range of operation. The propagation losses encountered at low and medium frequencies are discussed in a recent paper by the author [32].

Finally it is of interest to examine briefly the behavior of the system loss at short spacings ($d \ll \lambda$) when the antennas have dimensions which are comparable to the wavelength and thus large with respect to their spacing. For this analysis it will be more convenient to express (55) in the following form:

$$z_l = r'_b \left\{ \left(1 - \frac{r_m^2}{r'_a r'_b}\right)\left(1 + \frac{x_m^2}{r'_a r'_b}\right) \right\}^{1/2} - i x_b + i \frac{r_m x_m}{r'_a}. \qquad (59)$$

If it is assumed that the transmitting and receiving antennas are each dipoles (either electric or magnetic) with identical dimensions (comparable to the wavelength) and identical orientations, than as d approaches zero, r_m approaches the antenna self resistances $r'_a = r'_b$. Since the antennas are large, they may have small reactances and large resistances (e.g., half-wave dipoles in free space); in this case the load resistance r_l and the source resistance r_s will both be very small and, since ϕ is now also very small, the system loss will by (56) still be approximately equal to zero. In this case again the power delivered to the load is much larger than the power radiated. This unusual situation may be understood when it is noted that $i_b \cong -i_a$ so that the net power radiated from two such closely spaced antennas is approximately proportional to $(i_a + i_b)^2$ and thus extremely small. It is, of course, well known that it is difficult to radiate much power from a directional antenna consisting of a pair of closely spaced elements with the currents in the two elements out of phase.

7. Variations of the System Loss with Time

The instantaneous signal power available on long circuits involving transmission through the ionosphere or troposphere will vary with time (i.e., fade) due, in part, to phase interference between the components arriving along various transmission paths. Over short periods of time, during which transmission conditions may be regarded as constant, the instantaneous available power may, in some cases, be regarded as distributed in a Rayleigh distribution, to a good approximation. This distribution may be expressed

$$Q(p_a > y) = \exp(-y/\bar{p}_a), \tag{60}$$

where p_a is the instantaneous available power, $\bar{p}_a$ is the average power and Q is the probability that the instantaneous power exceeds some given value y. A discussion of the physical conditions under which the Rayleigh distribution may be expected to apply, together with an extension to the case in which a large nonfading component of power is present at the receiver terminals, is given in a recent paper by Norton, Vogler, Mansfield, and Short [9].

It has been found experimentally that $\bar{p}_a$ varies relatively slowly with time, and it is customary to determine the hourly median values p_h which, by (60), are related to $\bar{p}_a$ by $p_h = (\log_e 2)\bar{p}_a$. It has been determined experimentally that the hourly median values, p_h, for ionospheric or tropospheric propagation at a given time of day and a given season of the year are log-normally distributed [10, 11, 12]. It follows from this that the hourly median system loss is normally distributed with a standard deviation which may be expressed in decibels. For example, if the distribution of the 30 daily values of the hourly median transmission loss for 8 to 9 p.m. in June is studied, it is found that these values will appear to be a sample from a normal distribution of such values. Thus a reasonably satisfactory description of the expected distribution of the hourly median (or average) transmission loss for a specified time of day and season of the year can be given in terms of the two parameters, the median (of the hourly medians from day to day) and the standard deviation (of these hourly medians), both expressed in decibels.

8. Percentage of Time That a Desired Signal is Free of Interference

The effective range of a radio system is determined when the desired signal becomes so weak that it is obscured by the presence of noise or of other types of interference for too large a percentage of time. Consider first the effect of noise. This is most readily determined by means of a generalization [13] of Friis' definition [14] of the noise figure of a radio receiver. Consider the network of figure 4. It is convenient to compare the desired signal power with the noise power in network (a), i.e., in the loss-free receiving antenna. Let p_n denote the average external noise power and set

$$p_n \equiv f_a k_B t b \qquad \text{(watts)} \tag{61}$$

where k_B is Boltzmann's constant and is equal to 1.38×10^{-23}, t is the absolute temperature in degrees Kelvin, and b is the effective bandwidth in cycles per second as defined by Friis [14]. Thus (61) effectively defines the noise figure f_a of network (a). Network (c) has a noise figure $f_c = r'/r$ which is simply the loss factor of the receiving antenna circuit. Similarly the transmission line loss is equivalent to a noise figure f_t which is the ratio of the power in-and-out of the line. Finally we may

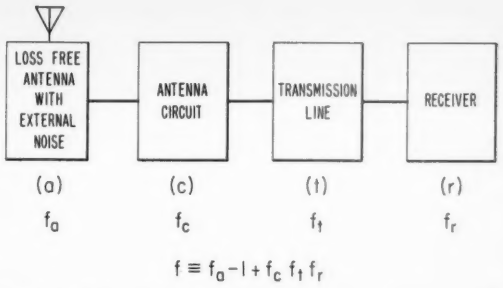


FIGURE 4. *Network defining the effective receiver noise figure including the effects of external noise.*

denote the noise figure of the receiver itself by f_r. Using Friis' method for combining the noise figures of several networks in tendem, we obtain for the effective noise figure at the input to the loss-free receiving antenna

$$f = f_a - 1 + f_c f_t f_r. \tag{62}$$

Finally, if we write r_n for the minimum signal-to-noise power ratio which will provide a given grade of reception, e.g., teletype reception with an 0.01 percent binary error rate, then the minimum signal power available at the terminals of the receiving antenna which will provide this grade of reception may be expressed

$$p_m = r_n f k_B t b \qquad \text{(watts)}. \tag{63}$$

Expressed in decibels (63) becomes

$$P_m = R_n + F + B - 204, \tag{64}$$

where $R_n = 10 \log_{10} r_n$, $F = 10 \log_{10} f$, $B = 10 \log_{10} b$, and $10 \log_{10} (k_B t) = -204$ if we use a reference temperature $t = 288.48°$ K.

The power P'_{ta} radiated from the transmitting antenna of a desired station at a which is required for a specified grade of reception by a receiver at b with a given transmission loss L_{ab} may now be expressed

$$P'_{ta} = L_{ab} + P_{mb} = L_{ab} + R_n + F_b + B - 204 \tag{65}$$

If the signal has a well-defined short term fading characteristic, e.g., representable by the distribution corresponding to that of a constant component plus a Rayleigh-distributed component [9], it is customary to generalize R_n to be the minimum median-signal to median-noise ratio required for a specified grade of reception of such a fading signal. In this case (65) gives the required radiated power if L_{ab} and F_b in (65) are considered to be hourly median values or, more generally, median values over a sufficiently short period of time that the short term variations of the signal and the noise each have the characteristics specified in determining R_n. In order to allow for the long-term fading characteristic, we recall that the hourly median values of the transmission loss at a specific time of day, season

of the year, and period in the sunspot cycle, tend to be normally distributed about their long-term median value L_{mab} with a standard deviation σ_{Lab}. Now let $\lambda(x)$ denote a standard normal deviate and x the probability that a randomly chosen normally distributed variable with unit standard deviation will not exceed $\lambda(x)$:

$$x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda(x)}\exp(-y^2/2)dy. \tag{66}$$

Note that $\lambda(x)$ varies from $-\infty$ to $+\infty$ as x varies from 0 to 1 and that $\lambda(0.5)=0$. Now we may express the value of the hourly median transmission loss not exceeded with probability x by

$$L_{ab}(x)=L_{mab}+\sigma_{Lab}\,\lambda(x). \tag{67}$$

Similarly F_b tends to be normally distributed about its long-term median value F_{mb} with a standard deviation σ_{Fb}:

$$F_b(x)=F_{mb}+\sigma_{Fb}\lambda(x). \tag{68}$$

Equations (67) and (68) are not valid in the limit as x approaches zero since L_{ab} and F_b are inherently positive, but this restriction is unimportant in practical applications. The random variables L_{ab} and F_b may be correlated (usually negatively) with correlation coefficient ρ_{abn} since diurnal, seasonal, and secular variations in propagation conditions sometimes cause L_{ab} to decrease at the same time that they cause F_b to increase. Now the above results may be combined and the following relation determined by standard statistical methods:

$$(\sigma_{Lab}^2+\sigma_{Fb}^2+2\rho_{abn}\sigma_{Lab}\sigma_{Fb})^{1/2}\lambda(0.01q_{abn})=P'_{ta}(q_{abn})-R_n-L_{mab}-F_{mb}-B+204. \tag{69}$$

In the above, q_{abn} is the percentage of hours during which the specified (or better) grade of reception of the signals from station a is possible at station b when noise is the only source of interference. Note that q_{abn} varies from 0 to 100 percent as the right hand side of (69) varies from $-\infty$ to $+\infty$. When $P'_{ta}(q_{abn})$ is fixed, (69) may be solved for q_{abn}; alternatively, by fixing q_{abn}, it is possible to solve (69) for the transmitter power $P'_{ta}(q_{abn})$ required to provide the specified (or better) grade of service for q_{abn} percent of the hours at a given time of day, season of the year, and period of the sunspot cycle. For example, if service is required for $q_{abn}=99$ percent of the hours, we find from tables of the standard normal deviate [15] that $\lambda(0.99)=2.326$ and thus $2.326(\sigma_{Lab}^2+\sigma_{Fb}^2+2\rho_{abn}\sigma_{Lab}\sigma_{Fb})^{1/2}$ db more power is required than would be required for $q_{abn}=50$ percent of the hours since $\lambda(0.5)=0$. If L_{mab}, σ_{Lab}, F_{mb}, σ_{Fb}, ρ_{abn}, and R_n are all known as a function of time, then $P'_{ta}(q_{abn})$ can be determined as a function of time.

For a proper determination of F_{mb} and σ_{Fb} it is necessary to know the magnitude and variance of the external noise f_a at the location b. Predictions of f_a over a wide range of frequencies and geogaphical locations were originally published by Crichlow et al. [16]. Revised predictions of f_a are available in a recent report [17] of the Consulting Committee on International Radio (C.C.I.R.). Using the results in a recent paper by Cottony and Johler [18], values of f_a may also readily be determined for cosmic radio noise simply by dividing their equivalent blackbody temperatures by 288.48. It should be noted that f_a will usually be somewhat dependent upon the receiving antenna directivity and the published values of f_a quoted above correspond to reception on a short vertical electric antenna in references [16] and [17] and to reception on a horizontal half-wave electric dipole in [18].

It is, of course, possible to determine the instantaneous distribution of the transmission loss by combining the short term and long term fading distributions, and a solution of this problem for the case of signals with a Rayleigh distribution over short periods of time and with a log-normal distribution of hourly medians has been given by Garner McCrossen [19]. However, it is usually more convenient to follow the procedure described above and absorb the effects of the short term fading distribution into the required signal-to-noise ratio.

To the degree that efficient use is made of the radio spectrum, interference from other radio stations rather than noise will often limit the effective range of the desired station. It is convenient in formulating this problem to consider initially the interference caused at receiving location b as a result of the operation of another system including a transmitter at c and a receiver at d. This special case will then be extended to determine the interference between additional pairs of stations and ultimately to a mobile or other broadcast type of system. If we write r_u for the minimum median-signal to median-interference power ratio at the terminals of the receiving antenna which will provide a given grade of reception of a desired signal with appropriate allowance for the modulation characteristics and the short term fading characteristics of both the desired and undesired carriers, then we find that the criterion for the absence of interference at the receiving location b from the undesired transmitter at c may be expressed

$$L_{cb}-L_{ab}+P'_{ta}-P'_{tc}\geq R_u. \tag{70}$$

In the above $R_u=10\log_{10}r_u$, L_{ab}, and L_{cb} denote the hourly median values of the transmission loss at the time in question for the paths ab and cb, respectively, while P'_{ta} and P'_{tc} are the powers radiated from the transmitting antennas at a and c, respectively. Now let L_{mcb} denote the long term median value of L_{cb}, σ_{Lcb} its standard deviation while ρ_{abcb} denotes the correlation coefficient between the variations of L_{ab} and L_{cb}. Now we may determine from the following equation the percentage of the hours q_{abc} at a given time of day, season of the year and phase of the sunspot cycle, during which reception

of a given grade *or better* is possible at station b of signals from station a on the assumption that the effects of all other sources of interference may be neglected except that from the station at c:

$$(\sigma_{ab}^2+\sigma_{cb}^2-2\rho_{abcb}\sigma_{ab}\sigma_{cb})^{1/2}\lambda(0.01q_{abc}) =L_{mcb}-L_{mab}+P'_{ta}-P'_{tc}-R_u. \quad (71)$$

For the efficient use of the radio spectrum it is most important to notice that the percentage of hours q_{abc} free of interference from the undesired station at c depends on the *ratio* $P'_{ta}-P'_{tc}$ of the desired and undesired station powers, whereas the percentage of hours q_{abn} free of interference from noise depends only on P'_{ta}.

An expression is also required for the percentage of hours q_{cda} of satisfactory reception of station c at station d on the assumption that the effects of all other sources of interference may be neglected except that from the station a:

$$(\sigma_{cd}^2+\sigma_{ad}^2-2\rho_{cdad}\sigma_{cd}\sigma_{ad})^{1/2}\lambda(0.01q_{cda}) =L_{mad}-L_{mcd}+P'_{tc}-P'_{ta}-R_u. \quad (72)$$

Now suppose that q_{abc} and q_{cda} are both sufficiently large (say greater than 99%) so that satisfactory service would be available in both systems except for the effects of noise, i.e., suppose that q_{abn} is less than 99 percent and that q_{cdn} is possibly, but not necessarily, also less than 99 percent. In this case we should, if this is feasible, increase P'_{ta} until $q_{abn}>99$ percent. However, if this is done, it may cause interference to the reception of station c by station d since q_{cda} will be decreased by an increase in P'_{ta} *unless* P'_{tc} is increased in the same proportion so that $P'_{tc}-P'_{ta}$ remains constant. The above argument is easily extended to any number of station pairs, and thus we have shown that a net improvement in the use of the spectrum for a given geographical configuration of radio stations can *always* be obtained by increasing the power of all of the stations by the same number of decibels until the power of each of the stations is adequate to reduce the level of the noise well below the level of the interference caused by the other stations. In practice there will be an economic limit above which such horizontal increases in station power are no longer feasible. When this limit is reached, then still further improvement of the use of the spectrum can *usually* be obtained by allocating additional stations or, in the case of a broadcast or mobile service, by reducing the spacing between the existing stations until the interference between them is comparable to the interference from noise.

The above analysis is, of course, only qualitative; thus, for example, some types of service might have to be available for much more than 99 percent of time. Nevertheless the principles of efficient allocation described above appear to be universally applicable and it is concluded that a satisfactory allocation of facilities to any portion of the spectrum can be considered to exist only when it is impossible to find a geographical location at which radio noise, rather than radio signals, may be observed. It also follows from the above discussion that an efficient allocation of broadcast stations necessarily implies the existence of large areas in which there will be mutual interference between the stations since otherwise it would be possible to observe radio noise, rather than radio signals, in these interference regions.

Equations (69), (71), and (72) are strictly applicable only when a single source of interference, identified by the last subscript on q, is present at the receiving location which is identified by the second subscript on q. A formal method of extending these results to the case where several sources of interference are present with comparable magnitudes at the receiving location was suggested in [1] and the use of this method for studying broadcast station allocations was described in a paper by Norton, Staras, and Blum [20]. However, because of the complexity of the above formal solution, it is seldom used in practice. Instead, it is usually more practical to consider that a receiving location is satisfactorily free of interference provided the values of q for all of these sources of interference are each independently larger than some predetermined value, say 99 percent.

It is of interest to see how the above interference analysis would be applied to the allocation of broadcast stations. In the first place it seems clear that the maximum economically feasible transmitter power and antenna height should be used so as to overcome the effects of noise for each station over the widest possible area. Having fixed the transmitter power and antenna height the stations should then be located near enough together so as to minimize the areas between the stations where noise may be observed free of radio signals. In fact, studies have been made [21] which indicate that there are optimum separations between broadcast stations with a fixed power; when the stations are located on a regular triangular lattice these optimum separations are quite large but, in the practical case of irregularly located stations, their optimum spacing will be much smaller and will be smallest in the case of isolated pairs of stations.

Variations of Path Antenna Gain With Time in Ionospheric Scatter Propagation

During recent years extensive measurements have been made at the Central Radio Propagation Laboratory on several long transmission paths in the United States, Canada, and Alaska of the system loss for the ionospheric scatter mode of propagation. Some of these measurements were reported in detail in a paper by Bailey, Bateman, and Kirby [22]. In this section a further analysis will be given of some of the data obtained by Bailey, Bateman, and Kirby on the Fargo-Churchill path.

The measurements were made in the following way. At the Fargo end of the path, the transmitter power was switched every half hour alternately to a high gain rhombic antenna and to a half-wave dipole antenna, both at the same height above the ground.

At the Churchill end of the path, continuous recordings were made of the signals from Fargo as received simultaneously on a high gain rhombic antenna and on a half-wave dipole antenna, both at the same height above the ground. For odd half hours, measurements were thus available of the half-hourly median system losses L_{rr} and L_{rh}, and for even half hours measurements were available of the median system losses L_{hr} and L_{hh}, where the first subscript denotes the transmitting antenna while the second subscript denotes the receiving antenna. The half-hourly median path antenna power gain corresponding to the use of a half-wave dipole antenna on the transmitting end of the path and a rhombic antenna on the receiving end of the path may now be determined from:

$$G_{pphr}=L_{hr}-L_{hh}+4.3. \tag{73}$$

In the above the term 4.3 is the assumed path antenna gain between the two loss-less half-wave dipoles. On the assumption that the system loss does not vary appreciably over a period of 1 hr, it is also possible to estimate the half-hourly median path antenna power gain corresponding to the use of rhombic antennas on both ends of the path from

$$G_{pprr}\cong L_{rr}-L_{hh}+4.3, \tag{74}$$

where L_{rr} and L_{hh} denote the values of these median system losses in successive half-hour periods of time. Only these latter values are analyzed here and in what follows we have set $G_{pprr}\equiv G_{pp}$, $L_{rr}\equiv L_s$ and $L_{hh}-4.3\equiv L_b$. The observed cumulative distributions of G_{pp}, L_s and L_b as thus defined are given on figure 5 for a 3-day period in February 1952. Note that each of these three observed quantities appears to be a sample from a normal distribution, and on this assumption we find the following sample statistics: $\overline{G}_{pp}=25.71$ db, $\sigma_{Gpp}=5.85$ db, $\overline{L}_s=176.75$ db, $\sigma_{Ls}=6.98$ db, $L_b=202.75$ db, and $\sigma_{Lb}=2.47$ db. The values of L_s and G_{pp} are highly correlated, the correlation coefficient determined from this sample being $\rho=-0.94$. On the assumption that L_s and G_{pp} are normally distributed, the random variable $L_b\equiv L_s+G_{pp}$ will also be normally distributed with expected mean $L_{be}=L_s+G_{pp}=202.46$ db and expected variance $\sigma^2_{Lbe}=\sigma^2_{Ls}+\sigma^2_{Gpp}+2\,\rho\,\sigma_{Ls}\,\sigma_{Gpp}=(2.48)^2$. Note that these values are in good agreement with the values of the mean and variance of L_b determined directly from the sample.

It is seen from the above analysis that the variance of the system loss for ionospheric scatter propagation increases with an increase in the path antenna gain. There is some evidence that this may not be the case for tropospheric scatter propagation. For example, Bullington, Inkster, and Durkee [33] measured the variance of 505–Mc and 4090–Mc tropospheric forward scatter over the same path using 28-ft parabolic reflectors in both cases, and thus with much higher gain on the higher frequency, and found an actually smaller variance for the 4090–Mc transmission losses.

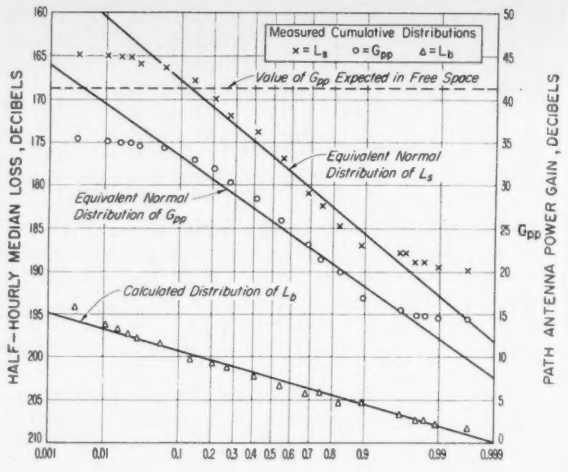


FIGURE 5. *Measurements of* L_s, L_b, *and* G_{pp} *for the Fargo-Churchill ionospheric forward scatter path on Feb. 15, 16, 17, 1952;* d=*824 statute miles,* f=*49.7 Mc.*

10. Transmission Loss for Tropospheric Forward Scatter and Ground Wave Propagation

As was pointed out in an earlier paper [1] the concept of transmission loss is especially appropriate for describing forward scatter propagation, and this idea was developed in some detail in a recent paper by Norton, Rice, and Vogler [12]. Recently the author showed [23,24] that the $(r/l_o)K_1(r/l_o)$ correlation function could be used for describing the physical nature of atmospheric turbulence and that its use, together with an assumed exponential dependence of the gradient of refractive index with height, provides a theory of tropospheric forward scattering in good agreement with the available data over a wide range of distances, frequencies, and antenna heights.

Recently, Bean and Thayer [25] developed several model tropospheric atmospheres which depend only upon the single parameter, N_s the value of the refractivity at the surface of the earth, i.e., $(n_s-1)\cdot 10^{-6}$ where n_s is the refractive index at the surface. In the present paper use will be made of the particular Bean and Thayer model in which N decreases linearly with height for the first kilometer above the surface, and exponentially with height above 1 km, having scale heights dependent on N_s for the height range from 1 to 9 km but having a fixed scale height above 9 km. Figure 6 compares the trace of radio rays in these model atmospheres for the typical values of N_s=250, 301, 350, and 400 and for several initial angles Θ_o of departure.

Using results of this kind, it is possible to use the methods of geometrical optics to calculate the groundwave fields expected in such atmospheres at points well within the horizon and to use the method described below for points beyond the horizon; it is then easy to join these curves and obtain a solutio

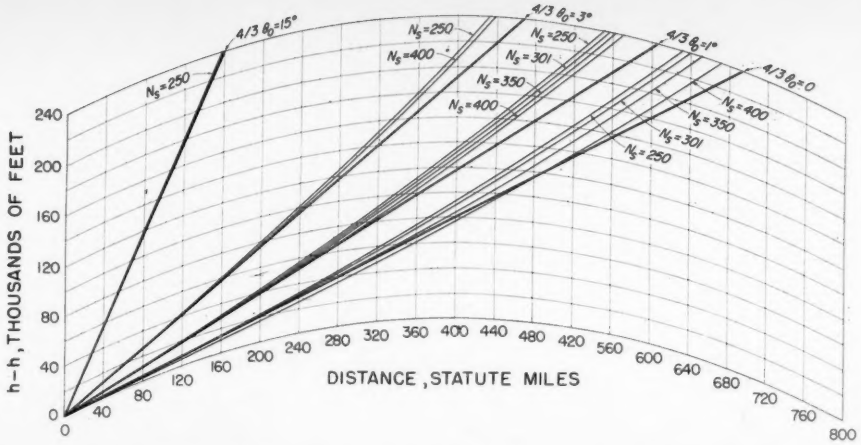


FIGURE 6. *Comparison of rays in the 1958 CRPL reference refractivity atmospheres with those for an earth with no atmosphere and an effective radius* $a_e = 5{,}280$ *statute miles* ($a_e/a = 4/3$).

for the groundwave fields expected for such atmospheres at any distance and for any antenna height. Furthermore, using the methods outlined in [23] and [24] and to be described more fully in a paper [26] to be published in this Journal in the near future, the tropospheric forward scattered fields may be calculated. Figures 7, 8, and 9 are examples of the median basic transmission loss for the groundwave and for the tropospheric scatter modes of propagation over a smooth spherical earth as calculated in this way. Allowance was made at the higher frequencies for atmospheric absorption by using the results of Bean and Abbott [27]. On figure 8 the first two oscillations of the field caused by interference between a direct and ground-reflected wave are shown for $d=10$ and 20 miles, but for the other distances only one oscillation is shown. The six points of field maxima are shown for all of the distances as circled points. Note that the total number of maxima to be expected (as a function of range for a given height or as a function of height at a given range) for a particular antenna height is equal to the number of half-wavelengths contained in this height; in the present example 30 ft contains 6 half-wavelengths at 100 Mc.

It is well known that approximate allowance may be made for the effects of radio refraction in the troposphere by calculating the fields for an earth with no atmosphere but with an effective radius a_e determined by

$$a_e = a \Big/ \left[1 + \frac{a}{n_s}\left(\frac{dn}{dh}\right)_s\right], \tag{75}$$

where a is the actual radius, n_s is the value of the refractive index at the surface and $(dn/dh)_s$ is the gradient (usually negative) of the refractive index at the surface. We will see that it is a simple matter to calculate the transmission loss L at or beyond the horizon for a horizontally homogeneous model atmosphere in terms of the transmission loss L_e expected for an earth with no atmosphere but with an effective radius a_e. Figure 10 illustrates the geometry of this more general method of making allowance for the effects of atmospheric refraction on the transmission loss between antennas at a and at b. If the index of refraction is specified as a function of height above the surface in such a way that ray tracing methods [25] may be used for establishing the trajectory of the ray, and such a restriction will rule out very few atmospheric models that are likely to be proposed, then it is possible to calculate for such atmospheres Δh_1 and Δd_1 as a function of the height h_1 and to calculate Δh_2 and Δd_2 as a function of the height h_2. Figures 11 and 12 show Δd and Δh versus h for the above-described Bean and Thayer model atmospheres for $N_s=250$, 301, 350, and 400, i.e., for $a_e=4878.50$, 5280, 5896.66, and 6996.67 statute miles, respectively. It may be noted that these particular model atmospheres have discontinuities in their gradients at 1 km above the surface and at 9 km above sea level, but these discontinuities cause no difficulties since we have used ray tracing methods based on geometrical optics. Note that the use of geometrical optics is preferable in this particular application; thus the use of a rigorous wave treatment of the problem would probably have lead to false conclusions since the actual atmosphere cannot have such abrupt discontinuities in its gradient.

If n were assumed to have a constant gradient $(dn/dh) = -n_s\left(\frac{1}{a} - \frac{1}{a_e}\right)$ at all heights (as determined by solving (75) for $(dn/dh)_s$ and then assuming that this surface gradient is the same at all heights), then it is easy to show for such an atmosphere that the distance to the radio horizon from a height h is given by

$$d = a_e \arccos\left[a_e/(a_e+h)\right] \cong \sqrt{2a_e h} \tag{76}$$

The approximation in the above is valid when h is

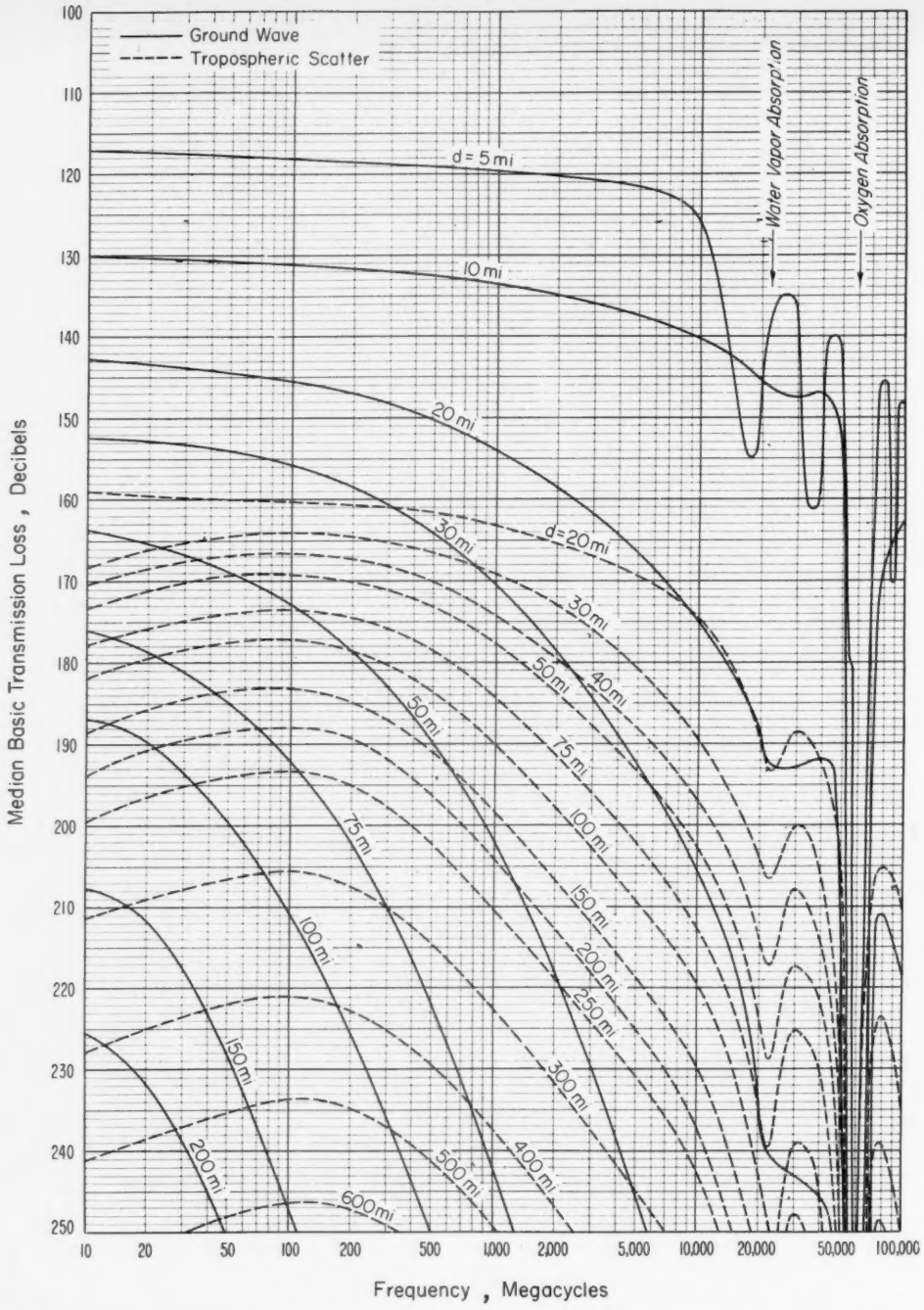


FIGURE 7. *Median basic transmission loss for the groundwave and tropospheric scatter modes of propagation over a smooth spherical earth; over land $\sigma=0.005$ mhos/m, $\epsilon=15$; horizontal polarization; transmitting and receiving antennas both 30 ft above the surface.*

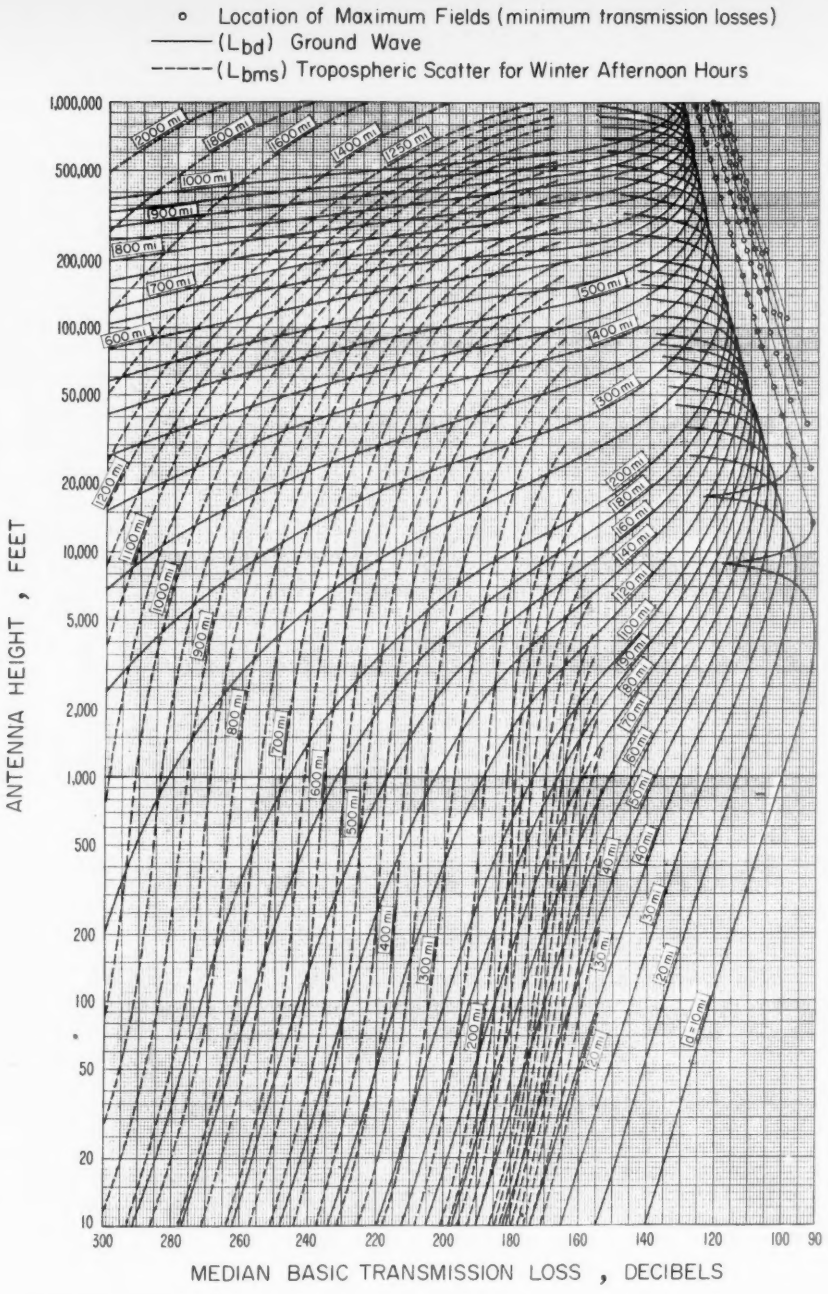


FIGURE 8. *Median basic transmission loss for the groundwave and tropospheric scatter modes of propagation at 100 Mc over a smooth earth in the 1958 CRPL reference refractivity atmosphere with* $N_s = 301$*; horizontal polarization over land; one antenna at 30 ft and the other antenna at the heights indicated.*

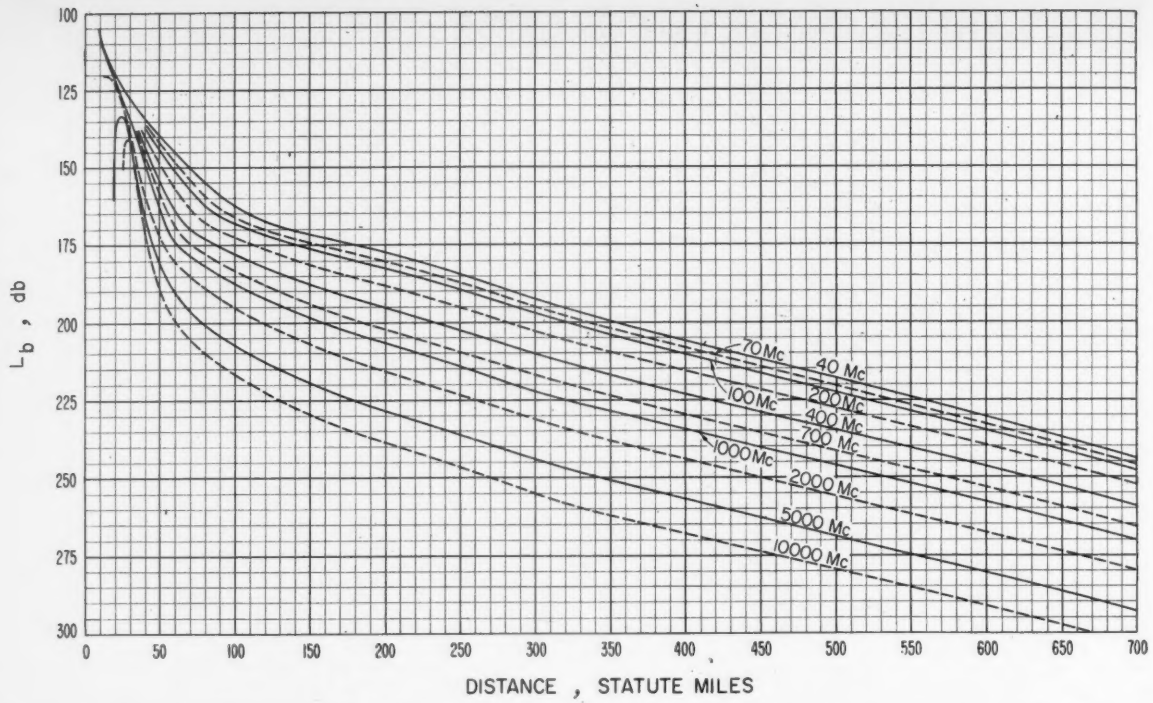


FIGURE 9. *Median basic transmission loss for the graundwave and tropospheric scatter modes of propagation over a smooth earth in the 1958 CRPL reference refractivity atmosphere with* $N_s=301$; *horizontal polarization over land; one antenna at 500 ft and the other antenna at 30 ft above the surface.*

sufficiently small. Since our model atmosphere has just such a linear gradient for the first kilometer above the surface, we see why Δh and Δd on figures 11 and 12 are equal to zero for heights h less than 1 km. At larger heights we see by figure 6 or by figures 11 and 12 how the distance to the radio horizon gradually departs from (76) in the model atmospheres which are much better approximations to the actual atmosphere than the assumption of a linear gradient at all heights.

Since the radio waves travel near the surface over that part of the path (of length d_s on fig. 10) between the radio horizons, it is intuitively clear that their propagation will be affected only by the refractive index gradient at the surface, and thus this portion of the propagation will be independent of the other characteristics of the atmosphere. Thus it is clear that only those portions of the propagation path from the antennas to their radio horizons are affected by that part of the atmosphere much above the surface, and these can be treated by the methods of geometrical optics. In this way the following relations are determined for evaluating the transmission loss $L(d,h_1,h_2)$ between the points a and b in the model atmosphere over an earth of radius a in terms of the transmission loss L_e expected for an earth with no atmosphere but with an effective radius a_e:

$$L(d, h_1, h_2)=L_e(d, h_1', h_2'), \qquad (77)$$

where

$$h_1'=h_1-\Delta h_1 \text{ and } h_2'=h_2-\Delta h_2$$

or

$$L(d, h_1, h_2)=L_e(d', h_1, h_2)-10\log_{10}(d'/d) \qquad (78)$$

where

$$d'=d+\Delta d_1+\Delta d_2$$

Equation (77) was used for determining the median basic transmission losses for the groundwave as shown on figure 8; essentially these same ray tracing methods were also used for determining the scatter losses shown on figures 7, 8, and 9, but in this case they were used also for tracing the rays up to the scattering region as well as for tracing the rays from an antenna to its radio horizon.

J. R. Wait, in a paper to be published shortly, has provided a rigorous basis for our intuitive development of (77) for sufficiently high frequencies and for model atmospheres in which the variation of the refractive index with height is monotonic.

Millington independently derived the above-described method of allowing for the effects of refraction in a horizontally homogeneous atmosphere, and suggested a formula for adjusting groundwave field strength similar to (78) but omitting the inverse distance factor $10\log_{10}(d'/d)$ which is usually negligible [28].

The groundwave curves in the C.C.I.R. Atlas [29] have been prepared using the approximately-linear

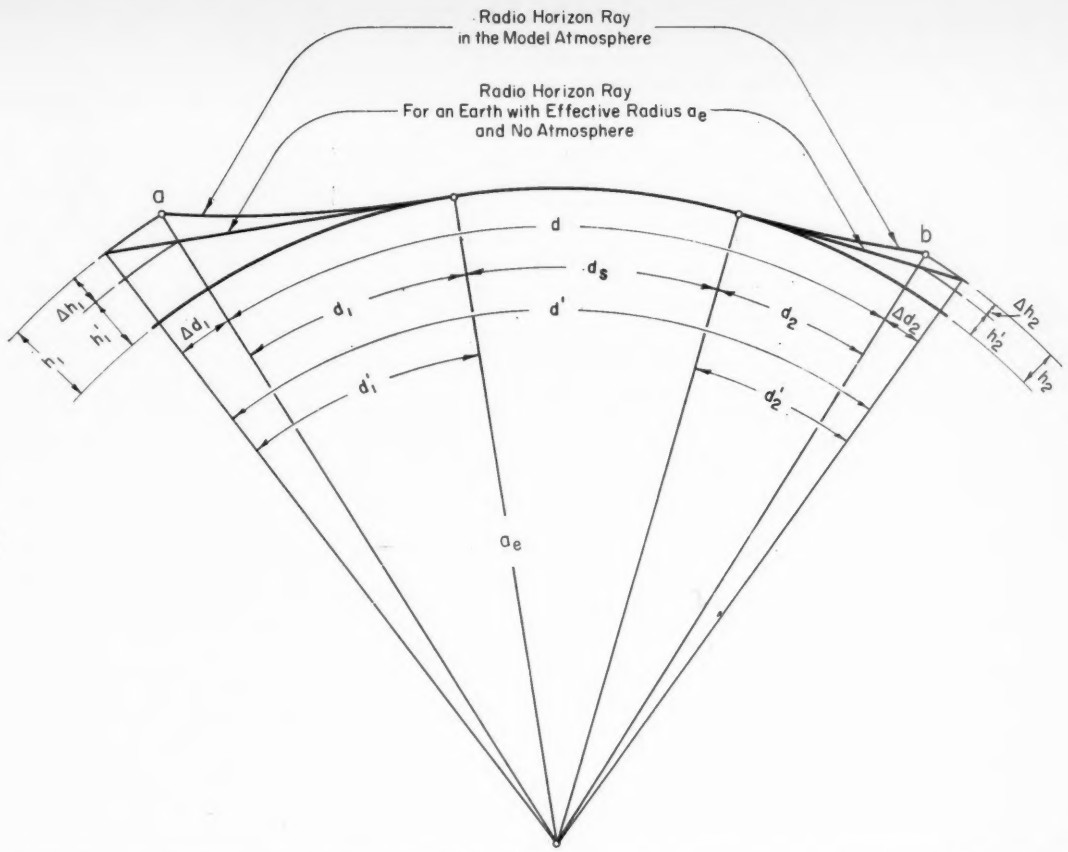


FIGURE 10. *Geometry illustrating the corrections* Δd *and* Δh *in the ray deviations between a model atmosphere and a sphere with no atmosphere but with an effective radius equal to* a_e.

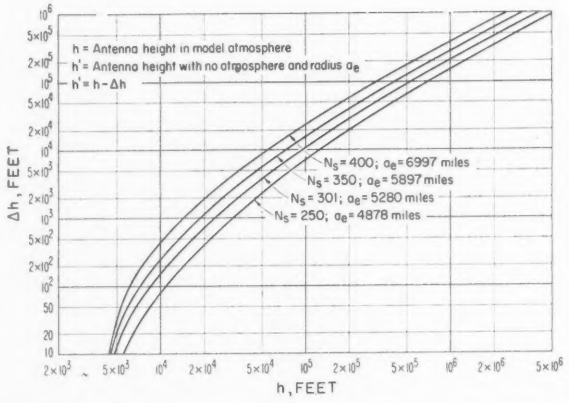

h = Antenna height in model atmosphere
h' = Antenna height with no atmosphere and radius a_e
h' = h − Δh
N_s = 400; a_e = 6997 miles
N_s = 350; a_e = 5897 miles
N_s = 301; a_e = 5280 miles
N_s = 250; a_e = 4878 miles
Δh, FEET
h, FEET

FIGURE 11. *Correction* Δh *to the heights of the radio horizon rays in the 1958 CRPL refractivity atmospheres relative to the heights they would have for spheres with no atmosphere but with effective radii* a_e.

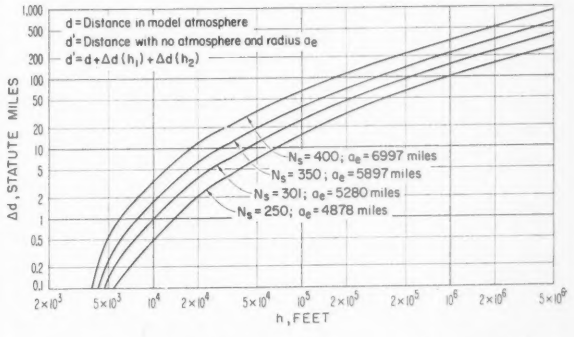

d = Distance in model atmosphere
d' = Distance with no atmosphere and radius a_e
d' = d + Δd (h1) + Δd (h2)
N_s = 400; a_e = 6997 miles
N_s = 350; a_e = 5897 miles
N_s = 301; a_e = 5280 miles
N_s = 250; a_e = 4878 miles
Δd, STATUTE MILES
h, FEET

FIGURE 12. *Corrections* Δd *to the distances to the radio horizon rays in the 1958 CRPL refractivity atmospheres relative to the distances they would have for spheres with no atmosphere but with effective radii* a_e.

Eckersley profile $n^2=1-\eta+\epsilon a^2/(a+h)^2$ for which the effective radius is given by $a_e=an_s^2/(1-\eta)$ and for which the distance to the radio horizon is given by $(a\,a_e)^{1/2}$ arc cos $[a/(a+h)]$. This horizon distance differs only slightly from that given by (76) for the atmosphere with a linear gradient, but it is necessary, for extremely high antennas, to use values of Δd and Δh obtained for an appropriate reference atmosphere if accurate results are to be expected. The groundwave curves in the present C.C.I.R. Atlas do not extend to sufficient heights to require much adjustment for nonlinearity of the atmosphere, but the curves in a recent atlas prepared for a linear atmosphere at the Radio Research Laboratories in Japan [30] extend to sufficient heights so that a small adjustment will be required. It is perhaps desirable to emphasize that these systematic corrections for atmospheric bending are not likely to be important in *most* practical applications, but it is aesthetically more satisfying and occasionally might be important to have a precise solution to this interesting problem.

Figures 7 and 8 give the values of median basic transmission loss separately for the diffracted groundwave and for tropospheric scatter. If we assume that the short term variations in the scatter fields are Rayleigh distributed, and that the groundwaves are relatively steady over short periods of time, then we may determine the expected combined median basic transmission loss L_{bm} in terms of the diffracted wave basic transmission loss L_{bd} and the median basic scatter transmission loss L_{bms} as follows:

$$L_{bm}=L_{bd}-R(0.5), \tag{79}$$

where $K=L_{bd}-L_{bms}+1.592$ is the ratio in decibels of the average scattered power to the diffracted wave power, and $R(0.5)$ is given graphically and in tables in [9]. When K is less than -16.5 db, L_{bm} differs from L_{bd} by less than 0.1 db, and when K is greater than 19.5 db, L_{bm} differs from L_{bms} by less than 0.1 db.

Finally, to determine the expected values $L_b(p)$ not exceeded by p percent of the hourly medians during a year, we may simply subtract $V(p,\theta)$ as given on figure 13 from L_{bm} as calculated from the values given on figures 7 or 8 or as given directly on figure 9. The parameter θ is the angular distance [12] and, over a smooth spherical earth, is given by

$$\theta=\frac{d_s}{a_e}=\frac{d-d_1-d_2}{a_e}, \tag{80}$$

where d_1 and d_2 are the distances to the horizons in the model atmosphere. The angular distance is a particularly convenient parameter for making appropriate allowance for the effects of irregularities in the terrain as is explained more fully in [12] and [26] and the curves on figure 13 may be used for a rough earth as well.

11. Range of Tropospheric Forward Scatter Systems

As an example of the method of using transmission loss in systems design, the problem of estimating the effective maximum range of a radio relay system using tropospheric scatter is considered. As an illustration of typical ranges to be expected, we will assume that the terrain is smooth, and will base our predictions on the above-described Bean and Thayer [25] model atmosphere with $N_s=301$. We will assume that either two 28-ft or two 60-ft parabolic antennas are used at both ends of the path, with their centers 30 ft above the ground and connected in a quadruple diversity system. With these assumptions, we may use the methods described in the preceding section to determine the transmission loss, $L(99)$, which we would expect one percent of the actual hourly median transmission losses to exceed throughout a period of 1 year; the use of these one percent losses implies that the specified service will be available for 99 percent of the hours. Tables 1 and 2 give for the 28-ft and 60-ft antennas the free space gains G_t+G_r, and the path antenna gains as a function of frequency and distance, while tables 3 and 4 give $L(99)$ as a function of frequency and distance.

The transmitter power P_t required to provide a specified type and grade of service for 99 percent of the hours may be obtained from the equation

$$P_t=P'_{ta}\ (99\%)+L_t=L_t+L(99)+R_n+F+B-204, \tag{79}$$

where L_t is the loss in the transmitting antenna circuit and this is set equal to 1 db in the following calculations. P'_{ta} (99%) may be determined from (69) and this is equal to the right-hand side of (79) since we will assume that the effective receiver noise figure has the constant value $F=5\ \log_{10}f_{Mc}-5$ so that $\sigma_{Fb}=0$ and $L(99)=L_{mab}+\sigma_{Lab}\ \lambda(0.99)$; it is assumed that the receiver incorporates gain adequate to ensure that the first circuit noise is detectable. $B\equiv 10\ \log_{10}(b_o+b_m)$ is the effective receiver bandwidth factor with b_o and b_m expressed in cycles per second; b_o allows for the drift between the transmitter and receiver oscillators, while b_m allows for the band occupied by the modulation.

For the calculations in this paper, the transmitter and receiver oscillators were each assumed to have a stability of one part in 10^8 and to vary independently so that $b_o=\sqrt{2}\ f_{Mc}\,10^{-2}$. Table 5 gives the values of b_m assumed for the various types of service considered. Table 5 also gives the values of R_n for the various kinds of service on the assumption that quadruble diversity is used. The value of R_n for the FM multichannel system is expected to provide a service with less than an 0.01 percent teletype character error rate. The FM multichannel system consists of 36 voice channels, each of which can accommodate sixteen 60 words per minute teletype circuits. The values of R_n given in table 5 were

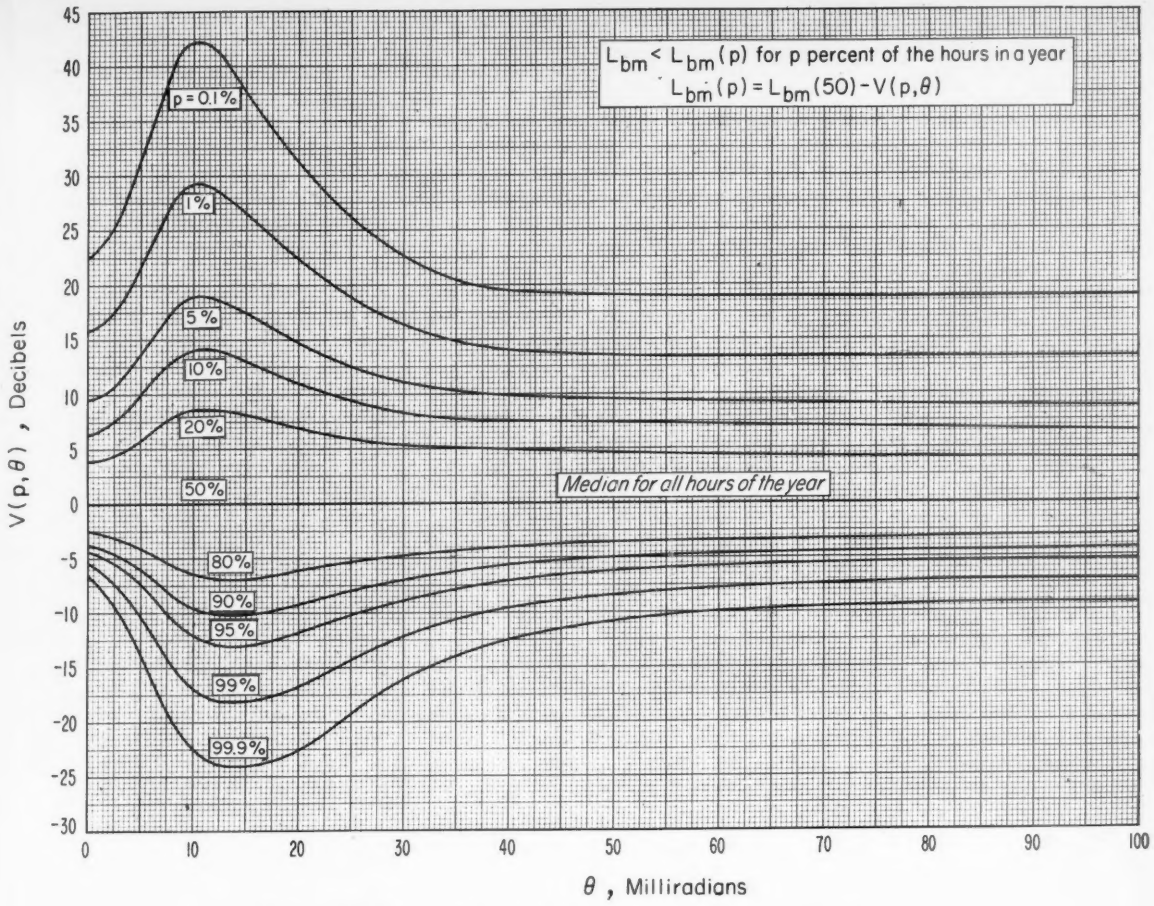


FIGURE 13. *Variance of transmission loss in tropospheric propagation.*

TABLE 1. *Path antenna gain for 28-ft parabolic antennas 30 ft above a smooth spherical earth with a CRPL model radio refractivity atmosphere corresponding to N_s=301*

M_0	G_t+G_r	G_p						
		d=100 mi	150	200	300	500	700	1,000
	db	db	db	db	db	db	db	db
100	33.02	33.02	33.02	33.02	32.92	32.82	32.72	32.67
150	40.07	40.07	39.97	39.97	39.87	39.57	39.47	39.37
200	45.06	45.06	44.96	44.86	44.66	44.36	44.16	44.06
300	52.11	52.03	51.91	51.71	51.31	50.81	50.51	50.41
500	60.98	60.75	60.38	60.08	59.38	58.48	58.18	58.18
700	66.83	66.35	65.83	65.23	64.33	63.23	62.83	62.93
1,000	73.02	72.12	71.22	70.42	69.22	67.92	67.32	67.22
1,500	80.07	78.32	76.57	75.87	74.27	72.57	71.67	71.57
2,000	85.06	82.48	80.66	79.26	77.16	75.46	74.36	73.86
3,000	92.11	87.71	84.21	83.41	80.91	78.51	77.61	77.31
5,000	100.98	93.28	90.18	87.68	84.68	82.18	81.28	80.98
7,000	106.83	96.53	92.83	90.03	86.83	84.12	83.33	82.93
10,000	113.02	99.32	95.22	92.52	89.02	86.02	85.22	84.92

TABLE 2. *Path antenna gain for 60-ft parabolic antennas 30 ft above a smooth spherical earth with a CRPL model radio refractivity atmosphere corresponding to N_s=301*

f_{M_0}	G_t+G_r	G_p						
		d=100 mi	150	200	300	500	700	1,000
	db	db	db	db	db	db	db	db
100	46.26	46.16	46.16	46.06	45.86	45.46	45.26	45.16
150	53.31	53.21	53.11	52.91	52.51	52.11	51.56	51.48
200	58.30	58.10	57.90	57.60	57.00	56.30	55.90	55.86
300	65.35	64.95	64.45	63.95	62.85	62.05	61.70	61.90
500	74.22	73.22	72.67	71.02	70.22	68.72	68.12	68.12
700	80.07	78.37	77.07	75.87	74.17	72.57	71.67	71.57
1,000	86.26	83.46	81.46	79.96	77.86	76.26	75.01	74.66
1,500	93.31	88.51	85.91	84.01	81.51	79.11	78.21	77.81
2,000	98.30	91.70	88.80	86.60	83.70	81.10	80.20	79.80
3,000	105.35	95.75	92.25	89.65	86.35	83.55	82.85	82.45
5,000	114.22	100.02	95.62	92.92	89.52	86.52	85.62	85.42
7,000	120.07	102.37	97.87	94.97	91.17	88.22	87.57	87.07
10,000	126.26	104.56	99.96	96.66	93.06	90.06	89.26	88.76

TABLE 3. *Transmission loss L(99) (corresponding to fields exceeded 99 percent of the time) expected between two 28-ft parabolic antennas at a height of 30 ft above a smooth spherical earth with a CRPL model radio refractivity atmosphere corresponding to $N_s=301$*

	$L(99)$ in decibels						
f_{Mc}	d mi						
	100	150	200	300	500	700	1,000
100	163	167	170	182	208	233	277
150	157	162	165	177	201	226	270
200	153	158	161	173	197	222	266
300	148	153	157	169	193	217	260
500	143	149	153	166	190	214	256
700	140	147	151	165	189	213	254
1,000	138	145	150	164	188	213	252
1,500	136	145	149	165	189	213	254
2,000	136	144	150	166	190	214	256
3,000	136	148	152	168	193	218	258
5,000	138	149	156	173	199	223	264
7,000	140	153	161	180	207	231	272
10,000	148	163	174	196	224	248	289

TABLE 4. *Transmission loss L(99) (corresponding to fields exceeded 99 percent of the time) expected between two 60-ft parabolic antennas at a height of 30 ft above a smooth spherical earth with a CRPL model radio refractivity atmosphere corresponding to $N_s=301$*

	$L(99)$ in decibels						
f_{Mc}	d=100 mi	150	200	300	500	700	1,000
100	150	154	157	169	195	221	264
150	144	149	152	164	188	214	257
200	140	145	148	161	185	210	254
300	135	141	144	158	182	206	249
500	131	137	142	155	180	204	246
700	127	136	140	155	179	204	245
1,000	127	135	140	156	180	205	244
1,500	126	136	141	157	182	207	248
2,000	126	136	142	159	184	209	250
3,000	128	139	145	163	188	212	253
5,000	131	143	151	169	195	219	260
7,000	134	148	156	176	203	227	268
10,000	142	158	170	192	219	244	285

determined by methods given in a recent report by Watt et al. [31]. The value of R_n for the FM multichannel system corresponds to typical fading encountered at 1,000 Mc, and this value of R_n may change by a few decibels with frequency as the fading changes, but such changes have so far not been evaluated quantitatively; furthermore, R_n will also change as the fading changes from hour to hour.

Table 6 gives as a function of frequency the maximum permissible hourly median transmission loss for a transmitter power of 10 kw: $L_M=204+P_t-L_t-R_n-F-B$ corresponding to the kinds of service described above. By combining the information in tables 3, 4, and 6, we can estimate the maximum range for a quadruple diversity system with 10-kw transmitters. These ranges are shown on figure 14 as a function of the radiofrequency.

TABLE 5

Type of service	b_m	R_n^a	Signal bandwidth	Post detection signal-to-noise ratio
	cps	*db*	*cps*	*db*
Transmission loss measurement.	0	0 [b]	0	
FM multichannel system.	3,750,000	9.5	36 voice channels each capable of use for sixteen 60 words per min teletype circuits.	0.01% teletype character error rate.
FM music	150,000	26.5	15,000	50.[c]
U.S. standard television.	3,750,000	32.7	3,750,000	30.[c]

[a] Ratio between the median intermediate frequency Rayleigh distributed signal and the rms Rayleigh distributed noise.
[b] Diversity reception not involved in this case.
[c] This ratio will be exceeded with a quadruple diversity system for 99 percent of each hour for which the corresponding value of R_n is maintained in each receiver.

TABLE 6. *Maximum permissible transmission loss L_M for 10 kw transmitters using the parameters of table 5*

f_{Mc}	Transmission loss measurement	FM multichannel	FM music	U.S. standard television
100	236.50	162.76	159.74	139.56
150	233.85	161.88	158.86	138.68
200	231.98	161.26	158.23	138.06
300	229.34	160.37	157.35	137.17
500	226.01	159.27	156.25	136.07
700	223.82	158.54	155.51	135.34
1,000	221.50	157.76	154.74	134.56
1,500	218.85	156.88	153.86	133.68
2,000	216.98	156.25	153.23	133.05
3,000	214.34	155.37	152.35	132.17
5,000	211.01	154.27	151.25	131.07
7,000	208.82	153.53	150.51	130.33
10,000	206.50	152.76	149.74	129.56

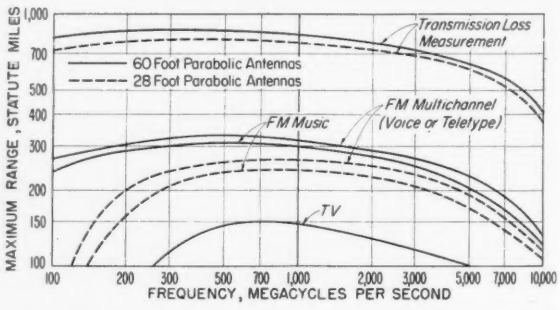


FIGURE 14. *Maximum distance at which satisfactory service of the types indicated may be provided for 99 percent of the hours using 10 kw. transmitters and quadruple diversity; smooth spherical earth; $N_s=301$; $h_t=h_r=30$ ft; atmospheric absorption and rain attenuation typical of the Washington, D.C., area.*

12. References

[1] Kenneth A. Norton, Transmission loss in radio propagation, Proc. IRE **41,** 146 (1953).

[2] S. A. Schelkunoff, Electromagnetic waves (D. Van Nostrand Co. 1943). See also S. A. Schelkunoff and Harald T. Friis, Antennas theory and practice (John Wiley & Sons, Inc., New York, N.Y., 1952).

[3] K. A. Norton, The propagation of radio waves over the surface of the earth and in the upper atmosphere, Pt. II, Proc. IRE **25,** 1203 (1937).

[4] James R. Wait, Radiation resistance of a small circular loop in the presence of a conducting ground, J. Appl. Phys. **24,** 646 (1953).

[5] James R. Wait and Walter J. Surtees, Impedance of a top-loaded antenna of arbitrary length over a circular grounded screen, J. Appl. Phys. **25,** 553 (1954).

[6] James R. Wait and W. A. Pope, Input resistance of L. F. unipole aerials, Wireless Engineer **32,** 131 (1955).

[7] James R. Wait, Effect of the ground screen on the field radiated from a monopole, IRE Trans. on Antennas and Propagation AP–**4,** 179 (1956).

[8] James R. Wait, Transmission of power in radio propagation, Electronic and Radio Engineer **36,** 146 (1959).

[9] K. A. Norton, L. E. Vogler, W. V. Mansfield, and P. J. Short, The probability distribution of the amplitude of a constant vector plus a Rayleigh-distributed vector, Proc. IRE **43,** 1354 (1955).

[10] Documents of the VIII Plenary Assembly of the C.C.I.R. (Warsaw, 1956), Published by the International Telecommunication Union, Geneva, 1957.

[11] P. L. Rice and F. T. Daniel, Radio transmission loss versus distance and antenna height at 100 Mc, Trans. IRE PGAP AP–**3,** 59 (1955).

[12] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** 1488 (1955).

[13] K. A. Norton and A. C. Omberg, The maximum range of a radar set, Proc. IRE **35,** 4 (1947).

[14] H. T. Friis, Noise figures of radio receivers, Proc. IRE **32,** 419 (1944).

[15] The Kelley Statistical Tables (Harvard University Press, Cambridge, Mass., 1948).

[16] W. Q. Crichlow, D. F. Smith, R. N. Morton and W. R. Corliss, Worldwide radio noise levels expected in the frequency band from 10 kilocycles to 100 megacycles, NBS Cir. 557 (Aug. 25, 1955).

[17] C.C.I.R. Report No. 65, Report on Revision of Atmospheric Radio Noise Data (Warsaw, 1956 available as a separate document from the International Telecommunications Union, Geneva, or in Vol. I of the Documents of the VIIIth Plenary Assembly of the C.C.I.R. Warsaw, 1956).

[18] H. V. Cottony and J. R. Johler, Cosmic radio noise intensities in the VHF band, Proc. IRE **40,** 1053 (1952).

[19] Garner McCrossen, Long-term cumulative distribution of instantaneous power (to be published).

[20] K. A. Norton, H. Staras and M. Blum, A statistical approach to the problem of multiple radio interference to FM and television service, Trans. IRE Professional Group on Antennas and Propagation PGAP–**1,** 43 (1952).

[21] Report of the Ad Hoc Committee for the Evaluation of the Radio Propagation Factors Concerning the Television and Frequency-Modulation Broadcasting Services in the Frequency Range Between 50 and 250 Mc, Federal Communications Commission Mimeo Nos. 36728; May 26, 1949; 36830; May 31, 1949; and 54382; July 7, 1950.

[22] D. K. Bailey, R. Bateman, and R. C. Kirby, Radio transmission at VHF by scattering and other processes in the lower ionosphere, Proc. IRE **43,** 1181 (1955).

[23] K. A. Norton, Point-to-point radio relaying via the scatter mode of tropospheric propagation, Transactions on Communications Systems **CS–4,** No. 1, 39 (March 1956).

[24] K. A. Norton, Recent experimental evidence favoring the $\rho K_1(\rho)$ correlation function for describing the turbulence of refractivity in the troposphere and stratosphere (to be published in the Proc. of the Joint Commission on Radio Meteorology held under the auspices of the International Scientific Radio Union at New York University, August 1957).

[25] B. R. Bean and G. D. Thayer, On models of the atmospheric radio refractive index, Proc. IRE **47,** 740 (1959).

[26] P. L. Rice, A. G. Longley, and K. A. Norton, Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss (to be submitted to this Journal).

[27] B. R. Bean and R. L. Abbott, Oxygen and water vapor absorption of radio waves in the atmosphere, Geophysica Pura E Applicata **37,** 127 (1957).

[28] G. Millington, Propagation at great heights in the atmosphere, Marconi Review **XXI,** 143 (4th Quarter 1958).

[29] Atlas of Ground-wave Propagation Curves for Frequencies between 30 Mc/s and 300 Mc/s (C.C.I.R. Resolution No. 11, published by the International Telecommunications Union, Geneva, 1955).

[30] Atlas of Radio Wave Propagation Curves for Frequencies between 30 and 10,000 Mc/s (prepared by the Radio Research Laboratories, Ministry of Postal Services, Tokyo, Japan, January 1958).

[31] A. D. Watt, E. F. Florman, and R. W. Plush, Carrier to noise requirements for teletype communications via non-fading or fading UHF carrier, J. Research NBS (in press).

[32] K. A. Norton, Low and medium frequency radio propagation, Proc. of the International Congress on the Propagation of Radio Waves at Liege, Belgium, October 1958 (to be published by the Academic Press).

[33] K. Bullington, W. J. Inkster, and A. L. Durkee, Results of propagation tests at 505 Mc and 4090 Mc on beyond-horizon paths, Proc. IRE **43,** 1306 (1955).

Boulder, Colo. (Paper 63D1–8).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Mode Expansion in the Low-Frequency Range for Propagation Through a Curved Stratified Atmosphere

H. Bremmer*

(January 8, 1959)

This expansion is particularly useful when considering ionospheric propagation at low frequencies. The complex problem dealing with two media, viz., a homogeneous earth and a surrounding stratified atmosphere, leads to intractable expressions. However, as the influence of the earth may be accounted for by an approximate boundary condition at the earth's surface, the problem is then reduced to that of the outer medium only. The coefficients of the mode expansion for this simplified problem will be derived while taking into account the earth's curvature; however, the latter proves to be negligible under very general conditions. The expansion to be derived is wanted in particular when studying the influence of a gradual transition in the electron density with height at the lower edge of the ionosphere.

1. Introduction

Maxwell's equations for time-harmonic functions proportional to exp $(-i\omega t)$ can be put in the following form (expressed in Gaussian units) provided the dielectric constant ϵ, the permeability μ and the conductivity σ are independent of time:

$$\left.\begin{aligned} \operatorname{curl}\vec{e}-i\frac{\omega}{c}\mu\vec{h}=0; \qquad \operatorname{curl}\vec{h}+i\frac{\omega}{c}\epsilon_c\vec{e}=0,\\ \operatorname{div}\left(\epsilon_c\vec{e}\right)=0; \qquad \operatorname{div}\left(\mu\vec{h}\right)=0; \end{aligned}\right\}, \tag{1}$$

the complex parameter $\epsilon_c=\epsilon+i4\pi\Sigma/\omega$ includes both dielectric and conductive properties.

We take $\mu=1$ and assume a stratified atmosphere of spherical symmetry, all properties of which can be considered as functions of the distance r to the center of the earth. We accordingly substitute $\epsilon_c=k^2(r)/k_0^2$, $k_0=\omega/c$ being the wave number in a vacuum; $k(r)$ can be interpreted as the local wave number at a height $r-a$ (a=earth's radius) above the surface of the earth. In the absence of the earth's magnetic field $\epsilon_c(r)$ constitutes a scalar. The Maxwell equations (1) can then be satisfied by the field

$$\vec{e}=\frac{i}{k^2(r)}\operatorname{curl}\operatorname{curl}\left\{k(r)\vec{Hr}\right\}; \quad \vec{h}=\frac{1}{k_0}\operatorname{curl}\left\{k(r)\vec{Hr}\right\}, \tag{2}$$

provided the Hertzian vector $H\vec{r}$ has a radial direction throughout space ($\vec{r}$ being a vector of length r in this direction); moreover, the amplitude H has to satisfy the following scalar wave equation in the outer space $r>a$:

$$(\Delta+k_{\text{eff}}^2)H=0, \tag{3}$$

in which the new parameter k_{eff}^2 is defined as follows [1][1]:

$$k_{\text{eff}}^2(r)=k^2(r)-k(r)\frac{d^2}{dr^2}\left\{\frac{1}{k(r)}\right\}. \tag{4}$$

The solution (2) is of the so-called electric type which includes the case of a vertical electric dipole to be considered hereafter. Let the dipole be placed in a homogeneous slab extending in the region $a<r<a'$ directly above the earth's surface, the upper boundary $r=a'$ of which may be identified with the lower edge of the ionosphere in the case of ionospheric propagation.

* Present address: Philips Research Laboratories, Eindhoven, The Netherlands.

[1] Figures in brackets indicate the literature references at the end of this paper.

Assuming a wave number k_0 inside this slab, the primary field of an infinitesimal vertical dipole can be derived from (2) by substituting for H the function:

$$H_{pr}(P)=\frac{Il}{cb}\frac{\exp\,(ik_0\,TP)}{TP}; \tag{5}$$

T and P here represent the positions of transmitter and receiver, TP their mutual distance, b the distance from T to the center of the earth, and Il the moment of the transmitter. The correctness of (5) is shown by comparing the associated radial Hertzian vector $H_{pr}\vec{r}$ with the conventional Hertzian vector for such a dipole. The latter Hertzian vector has a constant direction (parallel to the dipole axis) throughout space, its amplitude being given by

$$\frac{Il}{c}\frac{\exp\,(ik_0TP)}{TP}.$$

It can be verified that both Hertzian vectors only differ by a gradient, namely that of the quantity $iIl\,\exp(ik_0TP)/(bk_0c)$; however, such a gradient has no effect whatever on the primary field which can therefore also be derived from the radial Hertzian vector $H_{pr}\vec{r}$ according to (5).

2. Form of the Mode Expansion

The homogeneous wave equation (3) does not hold at the transmitter itself. The latter involves a singularity of the amplitude H at the point T. In view of (5) this singularity is the same as that of the function $Il/(cbTP)$. We introduce polar coordinates with the origin at the center of the earth and the axis $\theta=0$ drawn through the transmitter, the position of which therefore is given by $r=b$, $\theta=0$. The singularity in question can be accounted for by a right-hand side of (3) which only differs from zero at T. The complete wave equation in the outer space $r>a$ thus proves to be:

$$(\Delta+k_{em}^2)H=-4\,\frac{Il}{cb^3}\,\delta(b-r)\,\frac{\delta(\theta)}{\theta}. \tag{6}$$

This equation can be verified by integrating both sides over an infinitesimal sphere around T, while applying Gauss's integral theorem to the left-hand side, and the following properties of the impulse function to the right-hand side:

$$\int_0^{\theta_0}\frac{\delta(\theta)}{\theta}\sin\,\theta d\theta=\int_0^{\theta_0}\delta(\theta)d\theta=\frac{1}{2};\qquad\int_{b-\epsilon}^{b+\epsilon}\delta(b-r)dr=1.$$

We next assume that the influence of the earth may be accounted for by the boundary condition

$$\frac{\partial}{\partial r}\,(rH)=\gamma rH,\qquad(r=a) \tag{7}$$

which is the equivalent of the corresponding condition $\partial H/\partial z=\gamma H$ for a plane boundary. The quantity $-i\gamma/\{k_0\epsilon_c(a)\}$ can be interpreted as a surface admittance (ratio at grazing incidence of the electric field and the magnetic field, both at infinitesimal distances above and below the earth's surface). For $\epsilon_c(a)=1$ the parameter γ is given, for a vertical dipole, by

$$\gamma=-i\,\frac{k_0^2}{k_1^2}\,(k_1^2-k_0^2)^{1/2},$$

k_1 being the complex wave number inside the earth. The special case of a perfectly conducting earth corresponds to $\gamma=0$. The value of γ can also be chosen such as to account for a stratification in the ground [2].

The approximative boundary condition (7) has two advantages in the present problem, viz., (a) it reduces the original two-media problem to that of the outer medium only, (b) it guarantees the existence of a complete set of discrete orthogonal modes in the outer space $a<r<\infty$. These modes are characterized as solutions of the wave equation (3) that satisfy both the boundary condition (7) at $r=a$ and a proper radiation condition at $r=\infty$.

The general form of the modes is obtained by determining solutions of (3) with the aid of a separation of variables. Owing to the axial symmetry of the present problem around the axis $\theta=0$ (containing the transmitter) we only have to concern the polar coordinates r and θ. The expansion of the scalar H in terms of the modes in question then becomes of the type:

$$H=\sum_{l=-\infty}^{\infty} c_l f_{n_l}(r) P_{n_l}\{\cos(\pi-\theta)\}; \tag{8}$$

the height-gain functions $f_n(r)$ occurring here are solutions vanishing at infinity of the differential equation

$$\frac{d^2}{dr^2}(rf_n)+\left\{k_{\text{eff}}^2(r)-\frac{n(n+1)}{r^2}\right\}(rf_n)=0; \tag{9}$$

finally the discrete eigenvalues n_l (constituting the order of a Legendre function) are to be determined from the boundary condition

$$\frac{d}{dr}(rf_n)=\gamma r f_n \qquad (r=a). \tag{10}$$

3. A First Evaluation of the Coefficients of the Mode Expansion

The condition (10) guarantees the required boundary condition (7) for the complete solution (8). All Legendre functions in (8) become infinite along the whole axis $\theta=0$, n_l being not an integer; however, the proper choice of the coefficients c_l will reduce this singularity to the required singularity at the transmitter T only. In view of properties of Legendre functions the singularity of each term of (8) appears from the approximation

$$P_n\{\cos(\pi-\theta)\}\sim\frac{2}{\pi}\sin(n\pi)\log\frac{\theta}{2}=\frac{2}{\pi}\sin(n\pi)\log\frac{\rho}{2r}, \quad (\theta\to 0 \text{ or } \rho\to 0) \tag{11}$$

ρ being the distance from the point (r, θ) to the axis $\theta=0$. The approximation (11) can be proved by applying Tauber's theorem [3]

$$\operatorname*{Lim}_{p\to 0}\left\{pf(p)\right\}=\operatorname*{Lim}_{t\to\infty}\frac{h(t)}{t}$$

to the operational relation $h(t)\doteqdot f(p)$ for which [3, p. 243, eq (35)]

$$h(t)=P_n(2e^{-t}-1)U(t); \qquad f(p)=\frac{\Gamma(p)\Gamma(p+1)}{\Gamma(p+n+1)\Gamma(p-n)},$$

while substituting $t=-2\log\sin(\theta/2)$. In its turn the approximation (11) involves the three-dimensional equation:

$$(\Delta+k_{eff}^2)[f_n(r)P_n\{\cos(\pi-\theta)\}]=\frac{4}{\pi}\sin(\pi n)\frac{f_n(r)}{r^2}\frac{\delta(\theta)}{\theta}. \tag{12}$$

This equation can be proved along the same lines as (6), but by integrating over an infinitesimal cylinder around $\theta=0$ instead of an infinitesimal sphere around T; the proof is facilitated by replacing $\delta(\theta)/(r^2\theta)$ by the equivalent quantity $\delta(\rho)/\rho$.

A summation of (12) over all eigenvalues, after an initial multiplication by c_n, yields, in view of (8):

$$(\Delta+k_{eff}^2)\,H=\frac{4}{\pi}\frac{1}{r^2}\frac{\delta(\theta)}{\theta}\sum_{l=-\infty}^{\infty}c_l\sin(\pi n_l)f_{n_l}(r).$$

The introduction of both negative and positive modes avoids the occurrence of a deviation of c_0 from the general expression for c_l valid for $l\neq 0$. In fact, it is well known from the special case of a perfectly conducting earth and ionosphere that the addition of the l and $-l$ mode to a single one involves an exceptional value of c_0 (Neumann factor). Obviously, the right-hand side of the latter relation has to be identical with that of (6), so as to have:

$$\frac{1}{\pi r^2}\sum_{l=-\infty}^{\infty}c_l\sin(\pi n_l)f_{n_l}(r)=-\frac{Il}{cb^3}\,\delta(b-r).$$

A further multiplication by r^2 yields:

$$\frac{1}{\pi}\sum_{l=-\infty}^{\infty}c_l\sin(\pi n_l)f_{n_l}(r)=-\frac{Il}{cb}\,\delta(b-r); \tag{13}$$

The right-hand side of (13) can also be expanded in terms of the functions $f_{n_l}(r)$. In fact, the complete orthogonality of the latter over the interval $a<r<\infty$ involves the following expansion for an arbitrary function $\varphi(r)$:

$$\varphi(r)=\sum_{l=-\infty}^{\infty}\frac{\int_a^{\infty}\varphi(r)f_{n_l}(r)\,dr}{\int_a^{\infty}f_{n_l}^2(r)\,dr}\,f_{n_l}(r).$$

An application to $\varphi(r)=\delta(b-r)$ results in:

$$\delta(b-r)=\sum_{l=-\infty}^{\infty}\frac{f_{n_l}(b)f_{n_l}(r)}{\int_a^{\infty}f_{n_l}^2(r)\,dr}.$$

Hence, by equating the coefficient of $f_{n_l}(r)$ in both sides of (13), we can evaluate c_l. The substitution of the value thus found into (8) yields the final form of the mode expansion, viz.,

$$H=-\pi\frac{Il}{cb}\sum_{l=-\infty}^{\infty}\frac{f_{n_l}(b)f_{n_l}(r)}{\int_a^{\infty}f_{n_l}^2(r)\,dr}\,\frac{P_{n_l}\{\cos(\pi-\theta)\}}{(\sin\pi n_l)}. \tag{14}$$

4. Another Representation of the Mode Expansion

The coefficients of (14) only depend on the eigen functions f_{n_l} at the transmitter and receiver, apart from the integral in the denominator. Fortunately, the latter can be reduced as follows to a quantity referring to the earth's surface; hence, the behavior of f_{n_l} elsewhere needs not to be known explicitly. We apply the operator

$$\frac{\partial}{\partial n}(rf_n)\cdot 1-rf_n\cdot\frac{\partial}{\partial n}$$

to (9). The resulting expression can be represented in the form:

$$(2n+1)\,f_n^2=\frac{d}{dr}\left\{rf_n\cdot\frac{\partial^2}{\partial r\partial n}(rf_n)-\frac{d}{dr}(rf_n)\cdot\frac{\partial}{\partial n}(rf_n)\right\},$$

in which d/dr denotes a differentiation with respect to r at constant n. An integration of this equation over the interval $a<r<\infty$, taking into account the vanishing at infinity of the expression between braces, yields a relation which can be put in the form:

$$(2n+1)\int_a^\infty f_n^2(r)\,dr=-\left[r^2 f_n^2\frac{\partial}{\partial n}\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}\right]_{r=a}.$$

We bear in mind that $(d/dr)\ (rf_n)/(rf_n)$ at $r=a$ does in general depend on n, though this quantity assumes the value γ independent of n_l at the eigenvalues n_l of n [see (10)]. The integral in (14) can thus be evaluated; it leads to the following alternative form of the mode expansion for H:

$$H=\frac{\pi Il}{ca^2b}\sum_{l=-\infty}^{\infty}\frac{(2n_l+1)}{\left[\frac{\partial}{\partial n}\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}\right]_{r=a;\ n=n_l}}\frac{f_{n_l}(b)\,f_{n_l}(r)}{f_{n_l}^2(a)}\frac{P_{n_l}\{\cos(\pi-\theta)\}}{\sin(\pi n_l)}. \tag{15}$$

We now pass to the vertical component E_r of the electric field, the most important field component for the vertical dipole under consideration. The field representation (2), combined with the wave equation (3), implies the formula:

$$E_r=\frac{i}{k}\left(k_{eff}^2+\frac{\partial^2}{\partial r^2}\right)(rH).$$

The operator $k_{\text{eff}}^2+\partial^2/\partial r^2$ only affects the height-gain factors $f_{n_l}(r)$ of (15). The resulting quantity reads, in view of (9),

$$\left(k_{\text{eff}}^2+\frac{\partial^2}{\partial r^2}\right)(rf_{n_l})=\frac{n_l(n_l+1)}{r}f_n.$$

We thus obtain from (15) the following further mode expansion:

$$E_r=\frac{i\pi Il}{kca^2br}\sum_{l=-\infty}^{\infty}\frac{n_l(n_l+1)(2n_l+1)}{\left[\frac{\partial}{\partial n}\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}\right]_{r=a;\ n=n_l}}\frac{f_{n_l}(b)\,f_{n_l}(r)}{f_{n_l}^2(a)}\frac{P_{n_l}\{\cos(\pi-\theta)\}}{\sin(\pi n_l)} \tag{16}$$

5. Connection of the Mode Amplitude With the Atmospheric Reflection Coefficient

The dependence of the mode coefficients in (16) on the height-gain functions $f_{n_l}(r)$ can be replaced by that on a properly defined reflection coefficient. In order to discuss this coefficient we assume a homogeneous layer $a<r<a'$ with wave number $k=k_0$ above the earth's surface. In the case of ionospheric propagation this layer may be identified with the space between the earth and the ionosphere, its upper boundary $r=a'$ then representing the lower edge of the ionosphere (at a height $h=a'-a$ above the earth's surface). In the case of the troposphere the thickness h might finally approach zero; the wave number in the homogeneous slab then had better be taken equal to $k_0'\neq k_0$ in order to account for the small difference of the refractive index at the earth's surface from its vacuum value of unity. We shall restrict ourselves to the ionospheric case; it then is sufficiently accurate to identify the wave number in the interspace between the earth and the ionosphere with its vacuum value k_0.

In the homogeneous slab the general solution of (9) is well known, viz., a combination of two spherical Hankel functions:

$$h_n^{(1)}_{(2)}(k_0r)=\left(\frac{\pi}{2k_0r}\right)^{1/2}H^{(1)}_{(2)\,n+1/2}(k_0r).$$

The corresponding complete wave functions $h_n^{(1)}(k_0r)P_n\{\cos(\pi-\theta)\}$ represent a rising and a descending wave, respectively. A special mode will in general contain both these waves. Hence the height-gain function has the form:

$$f_n(r)=\alpha_n h_n^{(1)}(k_0r)+\beta_n h_n^{(2)}(k_0r) \qquad (a<r<a'). \tag{17}$$

Near the earth's surface the logarithmic derivative in the denominator of (16) therefore can be represented as follows:

$$\frac{\frac{d}{dr}(rf_n)}{rf_n}=\frac{\alpha_n\frac{d}{dr}\{rh_n^{(1)}(k_0r)\}+\beta_n\frac{d}{dr}\{rh_n^{(2)}(k_0r)\}}{r\{\alpha_n h_n^{(1)}(k_0r)+\beta_n h_n^{(2)}(k_0r)\}}. \tag{18}$$

This quantity depends, amongst others, on the ratio α_n/β_n which can be connected with a reflection coefficient. In fact, the second term of (17) constitutes the reflected wave that corresponds to the rising wave represented by the first term; hence we can interpret the ratio of both terms at $r=a'$ as an atmospheric reflection coefficient $T(n)$. We then obtain:

$$T(n)=\frac{\beta_n h_n^{(2)}(k_0a')}{\alpha_n h_n^{(1)}(k_0a')}, \tag{19}$$

with the aid of which (18) can be transformed as follows at $r=a$:

$$\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}_{r=a}=\frac{\frac{d}{dr}\{rh_n^{(1)}(k_0r)\}_{r=a}+T(n)\frac{h_n^{(1)}(k_0a')}{h_n^{(2)}(k_0a')}\cdot\frac{d}{dr}\{rh_n^{(2)}(k_0r)\}_{r=a}}{a\left\{h_n^{(1)}(k_0a)+T(n)\frac{h_n^{(1)}(k_0a')}{h_n^{(2)}(k_0a')}h_n^{(2)}(k_0a)\right\}}. \tag{20}$$

For a stratification degenerating to a homogeneous medium in the domain $r>a'$ the coefficient $T(n)$ proves to reduce to the Fresnel reflection coefficient for an elevation angle (compare next section) $\tau=\text{arc sin}\,(n/k_0a)$ if, moreover, the earth's curvature is neglected.

The rigorous expression (20) has to be differentiated with respect to n in order to obtain the denominator of (16). We shall derive a suitable approximation of this complicated quantity in the next section; it is then convenient to convert (20) into the alternative form:

$$\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}_{r=a}=\left[\frac{\frac{d}{dr}\{rh_n^{(1)}(k_0r)\}}{rh_n^{(1)}(k_0r)}\right]_{r=a}\frac{1+T(n)\frac{h_n^{(1)}(k_0a')}{h_n^{(2)}(k_0a')}\left[\frac{\frac{d}{dr}\{rh_n^{(2)}(k_0r)\}}{\frac{d}{dr}\{rh_n^{(1)}(k_0r)\}}\right]_{r=a}}{1+T(n)\frac{h_n^{(1)}(k_0a')}{h_n^{(2)}(k_0a')}\frac{h_n^{(2)}(k_0a)}{h_n^{(1)}(k_0a)}}. \tag{21}$$

6. An Approximation of the Mode Expansion (16)

The approximations for the function $h_n(z)$ have a different analytical form according to the position in the complex plane of the parameter z/n. We restrict ourselves to those values of z/n that correspond (for real $z\sim k_0a$) to eigenvalues n_l of n of type I, using Rydbeck's nomenclature [4]. These eigenvalues can be characterized by the unequalities:

$$\text{Re}\,n<k_0a-(k_0a)^{1/3},\ \text{Im}\,n>0,|n|\gg 1.$$

Such eigenvalues prove to play the dominating role in long-wave propagation. The approximations valid for the corresponding range of z/n read [5]:

$$h_n^{(1)}(z) \sim \frac{\exp\left\{(n+\tfrac{1}{2})\int_1^{z/n}(1-u^2)^{1/2}\frac{du}{u}\right\}}{(nz)^{1/2}\left(1-\frac{z^2}{n^2}\right)^{1/4}}; \quad h_n^{(2)}(z) \sim i\,\frac{\exp\left\{-(n+\tfrac{1}{2})\int_1^{z/n}(1-u^2)^{1/2}\frac{du}{u}\right\}}{(nz)^{1/2}\left(1-\frac{z^2}{n^2}\right)^{1/4}}, \tag{22}$$

in which the square roots are assumed as having a positive real part.

In view of the quantities occurring in (21) we are particularly interested in the two following further approximations which can be derived at once from (22):

$$\frac{h_n^{(1)}(z)}{h_n^{(2)}(z)} \sim -i\,\exp\left\{2(n+\tfrac{1}{2})\int_1^{z/n}(1-u^2)^{1/2}\frac{du}{u}\right\}, \tag{23}$$

$$\frac{\frac{d}{dz}\left\{z h_n^{(1)}(z)\right\}}{z h_n^{(1)}(z)} \sim \frac{1}{2z}+\frac{z}{2(n^2-z^2)} \pm \frac{(n^2-z^2)^{1/2}}{z}. \tag{24}$$

We next substitute $n=k_0 a \sin\tau$ or $n=k_0 a' \sin\tau'$. The asymptotic approximations of the complete wave functions $h_n^{(1)}(k_0 r)\,P_n\{\cos(\pi-\theta)\}$ then show that τ and τ' can be interpreted as the (in general complex) angles between the ray trajectories associated with the mode in question, and the verticals at $r=a$ and $r=a'$ respectively. Moreover, we substitute for z $k_0 a$ or $k_0 a'$. The corresponding transformations of (23) and (24) lead, amongst others, to the following approximations:

$$\frac{h_n^{(1)}(k_0 a')}{h_n^{(2)}(k_0 a')} \sim i\,\exp\left[2ik_0\left\{a'\cos\tau'-a\sin\tau\left(\frac{\pi}{2}-\tau'\right)\right\}\right],$$

$$\frac{h_n^{(2)}(k_0 a)}{h_n^{(1)}(k_0 a)} \sim i\,\exp\left[-2ik_0 a\left\{\cos\tau-\sin\tau\left(\frac{\pi}{2}-\tau\right)\right\}\right],$$

$$\left[\frac{\frac{d}{dr}\left\{r h_n^{(1)}(k_0 r)\right\}}{r h_n^{(1)}(k_0 r)}\right]_{r=a} \sim k_0\left(\mp i\cos\tau-\frac{\tan^2\tau}{2k_0 a}\right).$$

An application of the two latter relations yield:

$$\left[\frac{\frac{d}{dr}\{r h_n^{(2)}(k_0 r)\}}{\frac{d}{dr}\{r h_n^{(1)}(k_0 r)\}}\right]_{r=a} = \left[\frac{\frac{d}{dr}\{r h_n^{(2)}(k_0 r)\}}{r h_n^{(2)}(k_0 r)}\right]_{r=a} \times \left[\frac{\frac{d}{dr}\{r h_n^{(1)}(k_0 r)\}}{r h_n^{(1)}(k_0 r)}\right]_{r=a}^{-1} \times \frac{h_n^{(2)}(k_0 a)}{h_n^{(1)}(k_0 a)}$$

$$= \sim k_0\left(i\cos\tau-\frac{\tan^2\tau}{2k_0 a}\right)\times\frac{1}{k_0\left(-i\cos\tau-\frac{\tan^2\tau}{2k_0 a}\right)}\times\frac{1}{i\,\exp\left[-2ik_0 a\left\{\cos\tau-\sin\tau\left(\frac{\pi}{2}-\tau\right)\right\}\right]}.$$

The final substitutions of all these approximations into (21) yields:

$$\left\{\frac{\frac{d}{dr}(r f_n)}{r f_n}\right\}_{r=a} \sim -\left(ik_0\cos\tau+\frac{\tan^2\tau}{2a}\right)\frac{1-T(n)e^{2ik_0 h\cos\tau}\,\frac{1+i\tan^2\tau/(2k_0 a\cos\tau)}{1-i\tan^2\tau/(2k_0 a\cos\tau)}}{1+T(n)e^{2ik_0 h\cos\tau}}, \tag{25}$$

provided we neglect the small difference between τ and τ' (this difference vanishes altogether in the flat-earth approximation).

In (16) we need the derivative of (25) with respect to n. According to the relation $n = k_0 a \sin\tau$ we have

$$\frac{\partial}{\partial n} = \frac{1}{k_0 a \cos\tau}\frac{\partial}{\partial \tau}.$$

A corresponding evaluation of the n-derivative of (25) results in:

$$\frac{\partial}{\partial n}\left\{\frac{\frac{d}{dr}(rf_n)}{rf_n}\right\}_{r=a} = \sim \frac{1}{a\cos\tau}\left\{i\sin\tau - \frac{\sin\tau}{k_0 a\cos^3\tau} - \left(i\cos\tau + \frac{\tan^2\tau}{2k_0 a}\right)\frac{\partial}{\partial\tau}\right\}$$

$$\frac{1 - T(n)e^{2ik_0 h\cos\tau}\dfrac{1 + i\tan^2\tau/(2k_0 a\cos\tau)}{1 - i\tan^2\tau/(2k_0 a\cos\tau)}}{1 + T(n)e^{2ik_0 h\cos\tau}}. \qquad (26)$$

Moreover, we shall apply the well-known asymptotic approximation for Legendre functions having a large complex order n with positive imaginary part, viz.,

$$\frac{P_n\{\cos(\pi-\theta)\}}{\sin(n\pi)} \sim -\left(\frac{2i}{\pi n\sin\theta}\right)^{1/2} e^{i(n+1/2)\theta}. \qquad (27)$$

A final substitution of both (25) and (27) into (16), while approximating n_l+1 and $n_l+1/2$ by $n_l = k_0 a \sin\tau_l$, yields the following expansion in the simplest case of a transmitter and receiver on the ground ($r=b=a$):

$$E_r \sim -2\frac{k_0^{3/2} Il}{c}\left(\frac{2i\pi}{a\sin\theta}\right)^{1/2} e^{i\theta/2}$$

$$\sum_{l=-\infty}^{\infty} \frac{\sin^{5/2}\tau_l\, e^{ik_0 a\theta\sin\tau_l}}{\left\{\tan\tau + \dfrac{i\tan\tau}{k_0 a\cos^3\tau} - \left(1 - \dfrac{i\sin^2\tau}{2k_0 a\cos^3\tau}\right)\dfrac{\partial}{\partial\tau}\right\}\left\{\dfrac{1 - T(n)e^{2ik_0 h\cos\tau}\dfrac{1 + i\sin^2\tau/2k_0 a\cos^3\tau}{1 - i\sin^2\tau/2k_0 a\cos^3\tau}}{1 + T(n)e^{2ik_0 h\cos\tau}}\right\}_{\tau=\tau_l}} \qquad (28)$$

7. Flat Earth Approximation of the Mode Expansion

The following simplification results from (28) for $a=\infty$, remembering the relation $\theta = d/a$ (d=distance from transmitter to receiver):

$$E_r \sim -2\frac{k_0^{3/2} Il}{c}\left(\frac{2\pi i}{d}\right)^{1/2} \sum_{l=-\infty}^{\infty} \frac{\sin^{5/2}\tau_l\, e^{ik_0 d\sin\tau_l}}{\left[\left(\tan\tau - \dfrac{\partial}{\partial\tau}\right)\left\{\dfrac{1 - T(n)e^{2ik_0 h\cos\tau}}{1 + T(n)e^{2ik_0 h\cos\tau}}\right\}\right]_{\tau=\tau_l}}. \qquad (29)$$

This expression can also be derived straightforwardly by applying the flat earth conditions right from the beginning, while replacing the boundary condition (7) at the earth's surface by

$$\frac{\partial\Pi}{\partial z} = \gamma\Pi \qquad (z=0). \qquad (30)$$

The rigorous expansion for the amplitude of the vertically directed Hertzian vector Π [to be substituted in (2) instead of $H\vec{r}$] then arrived at reads:

$$\Pi = \frac{2\pi i}{c} Il \sum_{l=-\infty}^{\infty} \frac{\lambda_l}{\left[\dfrac{\partial}{\partial\lambda}\left\{\dfrac{df_\lambda/dz}{f_\lambda}\right\}\right]_{z=0;\,\lambda=\lambda_l}} \frac{f_{\lambda_l}(h_0) f_{\lambda_l}(z)}{f_{\lambda_l}^2(0)} H_0^{(1)}(\lambda_l d). \qquad (31)$$

In this series, replacing (15), $f_\lambda(z)$ denotes a height-gain function solving the differential equation

$$\frac{d^2 f_\lambda}{dz^2}+\{k_{\text{eff}}^2(z)-\lambda^2\}f_\lambda=0,$$

and satisfying the radiation condition at infinity ($z=\infty$). The other boundary condition (30) involves the relation

$$f_\lambda'(0)=\gamma f_\lambda(0) \qquad (z=0) \tag{32}$$

at the earth's surface [replacing (7)]; this condition can only be satisfied if λ equals one of the eigenvalues λ_l; h_0 represents the height of the transmitter above the earth. The expansion for the vertical component of the electric field [replacing (16)], to be derived from (31), becomes:

$$e_z=-\frac{2\pi}{k_0 c} Il \sum_{l=-\infty}^{\infty} \frac{\lambda_l^3}{\left[\frac{\partial}{\partial\lambda}\left\{\frac{df_\lambda/dz}{f_\lambda}\right\}\right]_{z=0;\,\lambda=\lambda_l}} \frac{f_{\lambda_l}(h_0)f_{\lambda_l}(z)}{f_{\lambda_l}^2(0)} H_0^{(1)}(\lambda_l d). \tag{33}$$

The wave functions in the homogeneous slab $o<z<h$ above the earth here reduce to simple exponentials so that (17) is to be replaced by:

$$f_\lambda(z)=\alpha_\lambda \exp\{-(\lambda^2-k_0^2)^{1/2}z\}+\beta_\lambda \exp\{(\lambda^2-k_0^2)^{1/2}z\}, \tag{34}$$

which enables an easy evaluation of the denominator in (29). The final series representing e_z for a transmitter and receiver on the ground ($h_o=z=0$) then proves to be, substituting moreover $\lambda_l=k_o \sin\tau_l$,

$$e_z=-\frac{2\pi i k_0^2}{c} Il \sum_{l=-\infty}^{\infty} \frac{\sin^3\tau_l}{\left[\left(\tan\tau-\frac{\partial}{\partial\tau}\right)\left\{\frac{1-T(\tau)e^{2ik_0h\cos\tau}}{1+T(\tau)e^{2ik_0h\cos\tau}}\right\}\right]_{\tau=\tau_l}} H_0^{(1)}(k_0 d\sin\tau_l). \tag{35}$$

This rigorous series leads to (29) if the Hankel function is replaced by its asymptotic approximation. A comparison of (29) and (28) shows under what circumstances the earth's curvature becomes negligible.

An evaluation of the derivative in the denominator of (35) leads to the alternative form:

$$e_z=-\frac{2\pi i k_0^2}{c} Il \sum_{l=-\infty}^{\infty} \frac{\sin^3\tau_l \quad H_0^{(1)}(k_0 d\sin\tau_l)}{\left[\tan\tau\,\frac{1-Te^{2ik_0h\cos\tau}}{1+Te^{2ik_0h\cos\tau}}-2\,\frac{e^{2ik_0h\cos\tau}(2ik_0h\sin\tau T-\partial T/\partial\tau)}{(1+Te^{2ik_0h\cos\tau})^2}\right]_{\tau=\tau_l}} \tag{36}$$

The denominator of (36) can further be worked out by considering the equation determining the modes. The latter is arrived at as follows. A substitution of (34) while evaluating the boundary condition (32) leads to the relation

$$\frac{(\beta/\alpha)-1}{(\beta/\alpha)+1}(\lambda^2-k_0^2)^{1/2}=\gamma \text{ for } \lambda=\lambda_l.$$

On the other hand, the definition of the atmospheric reflection coefficient for the flat case involves the following analogy of (19):

$$T(\lambda)=\frac{\beta}{\alpha} e^{2h(\lambda^2-k_0^2)^{1/2}}.$$

An elimination of β/α from the two latter relations yields an equation which can be put in the following form [replacing λ by $k_0 \sin\tau$ and $(\lambda^2-k_0^2)^{1/2}$ by $-ik_0\cos\tau$]:

$$\frac{ik_0\cos\tau_l+\gamma}{ik_0\cos\tau_l-\gamma}\cdot T(\tau_l)e^{2ik_0h\cos\tau_l}=1.$$

The first quotient here proves to be nothing else but the approximative plane-wave reflection coefficient at the earth's surface, $R(\tau_l)$ say, that results from the boundary condition (32). Hence the mode equation reads:

$$R(\tau_l)\,T(\tau_l)e^{2ik_0h\cos\tau_l}=1. \tag{37}$$

This simple equation is also obtained at once from the following resonance condition. We consider a plane rising wave in the homogeneous slab $0<z<h$, and derive the other plane rising wave that results from it after a reflection against the lower edge of the stratification and a reflection thereafter against the earth surface. The relation (37) then expresses the condition that the latter rising wave should be identical to the original one.

The exponential $\exp(2ik_0h\cos\tau_l)$ in (36) can now be replaced throughout by $\{R(\tau_l)\ T(\tau_l)\}^{-1}$. This yields the following expansion replacing (36):

$$e_z=\frac{2\pi ik_0^2}{c}\,Il\sum_{l=-\infty}^{\infty}\frac{\{1+R(\tau_l)\}^2\ \sin^3\tau_l}{\{(1-R^2)\tan\tau+2R(2ik_0h\sin\tau-\partial\log T/\partial\tau\}_{\tau=\tau_l}}H_0^{(1)}(k_0d\sin\tau_l). \tag{38}$$

For completeness sake we also mention the corresponding asymptotic approximation based on large values of k_0d, viz.,

$$e_z\sim\frac{(\pi i)^{1/2}(2k_0)^{3/2}}{c}\,\frac{Il}{d^{1/2}}\sum_{l=-\infty}^{\infty}\frac{\{1+R(\tau_l)\}^2\sin^{3/2}\tau_l\cos\tau_l}{\left\{1-R^2+2R\cos\tau\left(2ik_0h-\dfrac{\partial\log T/\partial\tau}{\sin\tau}\right)\right\}_{\tau=\tau_l}}e^{ik_0d\sin\tau_l}.$$

8. Special Case of a Flat Stratification Reducing to a Homogeneous Atmosphere

A further reduction of (38) is only possible by introducing a special model for the stratification in the space $z>h$. Let us consider the example of a homogeneous atmosphere, sharply bounded at $z=h$. The other boundary $z=0$ of the interspace $0<z<h$ also being sharp, the situation becomes almost identical to that of a wave guide limited by two infinite plane parallel plates with a separation h. However, we have to bear in mind the approximative boundary condition applied at $z=0$ so that this wall behaves differently from the other wall at $z=h$. Let us mark the refractive index of the atmosphere by n_i; the reflection coefficient T then reduces to an expression of the Fresnel type, viz.,

$$T(\tau)=\frac{n_i^2\cos\tau-(n_i^2-\sin^2\tau)^{1/2}}{n_i^2\cos\tau+(n_i^2-\sin^2\tau)^{1/2}}.$$

This formula involves the relation

$$\frac{\partial\log T}{\partial\tau}=\frac{2\tan\tau(1+T)}{(1-T)(\sin^2\tau-n_i^2\cos^2\tau)},$$

with the aid of which (38) can further be transformed [6] into:

$$e_z=\frac{2\pi ik_0^2}{c}\,Il\sum_{l=-\infty}^{\infty}\left[\frac{(1+R)^2\sin^2\tau\cos\tau}{1-R^2+4R\left\{ik_0h\cos\tau-\dfrac{1+T}{(1-T)\ (\sin^2\tau-n^2\cos^2\tau)}\right\}}\right]_{\tau=\tau_l}H_0^{(1)}(k_0d\sin\tau_l). \tag{40}$$

Owing to the approximative boundary condition at $z=0$ we here obtain a completely discrete sum. The rigorous boundary condition at $z=0$, on the contrary, would yield also an integral contribution which can be ascribed to a continuous part of the mode spectrum. However, the latter vanishes in the case of a perfectly conducting wall at $z=0$.

The result (40) is also rigorous for a wave guide limited by two perfectly conducting walls at $z=0$ and $z=h$; we then have $T=R=1$. In this case we derive from (40) the following series for the vertical field component at $z=0$ that is due to a dipole also situated at $z=0$:

$$e_z=\frac{2\pi k_0}{ch} Il \sum_{l=-\infty}^{\infty} \sin^2\tau_l \, H_0^{(1)} (k_0 d \sin \tau_l).$$

The corresponding asymptotic approximation becomes:

$$e_z \sim 2^{3/2}\left(\frac{\pi k_0}{id}\right)^{1/2} \frac{Il}{ch} \sum_{l=-\infty}^{\infty} \sin^{3/2}\tau_l \, H_0^{(1)} (k_0 d \sin \tau_l);$$

moreover, we may substitute $\tau_l=\text{arc cos } (\pi l/k_0 h)$.

This work was done at the United States National Bureau of Standards, while the author was on leave from Philips Research Laboratories (Eindhoven, Netherlands). He is indebted to the Radio Propagation Engineering Division of the National Bureau of Standards for providing him with the opportunity to work in Boulder.

9. References and Notes

[1] H. Bremmer, Terrestial radio waves, p. 139 (New York, 1949).
[2] J. R. Wait, Radiation from a vertical antenna over a curved stratified ground, J. Research NBS **56,** 237 (1956) RP2671.
[3] Balth. Van der Pol and H. Bremmer, Operational calculus, p. 129, eq (15) (Cambridge, 1955).
[4] O. E. H. Rydbeck, On the propagation of radio waves, p. 72 (Goteborg, 1944).
[5] See reference 1, p. 38, figure 6, case II for functions of the first kind and case I for functions of the second kind; equivalent expressions are given in Watson, Theory of Bessel functions, p. 234 (Cambridge, 1941).
[6] The convergence and numerical treatment of series such as eq (40) have been discussed by J. R. Wait, On the mode theory of vlf ionospheric propagation, Geofis. pura e appli. **37,** 103 (1957).

Boulder, Colo. (Paper 63D1–9).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Transmission and Reflection by a Parallel Wire Grid

Martin T. Decker

(February 19, 1959)

A comparison is made at X-band frequencies of the theoretical and measured transmission and reflection coefficients for a parallel wire grid. The methods used are applicable to the measurements of these factors for various building materials.

1. Introduction

As a preliminary phase in the measurement of the attenuating properties of various inhomogeneous materials and structures, tests have been made with an obstacle having electrical behavior which lends itself to calculation. These data will then be useful in evaluating the validity of certain test configurations. A planar grid of parallel wires was chosen as a suitable obstacle. The transmission and reflection of electromagnetic waves by such a grid have been treated in various papers, and solutions have been presented for both oblique angles of incidence and arbitrary polarization of plane waves.[1,2] A comparison of one of these solutions with experiment has been made at frequencies near 9 kMc.

2. Calculation of Grid Transmission and Reflection

Reflection from a parallel wire grid may be calculated according to the method given by Wait.[1] The coordinate system used is shown in figure 1. Wait's paper treats the grid as infinite in the y-z plane, with the incident energy in the form of a plane wave with arbitrary polarization and angle of incidence. A special case is taken here in which the incident wavefront is a vertical plane, vertically polarized (electric vector parallel to the grid wires) and is incident on the grid at some angle θ measured in the x-y plane.

In this special case, the reflection coefficient is given by

$$R=\frac{-1}{1+2Z_g/K_0},$$

where Z_g is the equivalent shunt impedance of the grid, K_g is the intrinsic impedance of the incident wave in the direction normal to the grid and the ratio is given by

$$Z_g/K_0=i\left(\frac{d}{\lambda}\right)\cos\theta\left[\ln\frac{d}{2\pi\alpha}+F\left(\frac{d}{\lambda},\theta\right)\right]+\cos\theta Z_i\frac{d}{\eta_0},$$

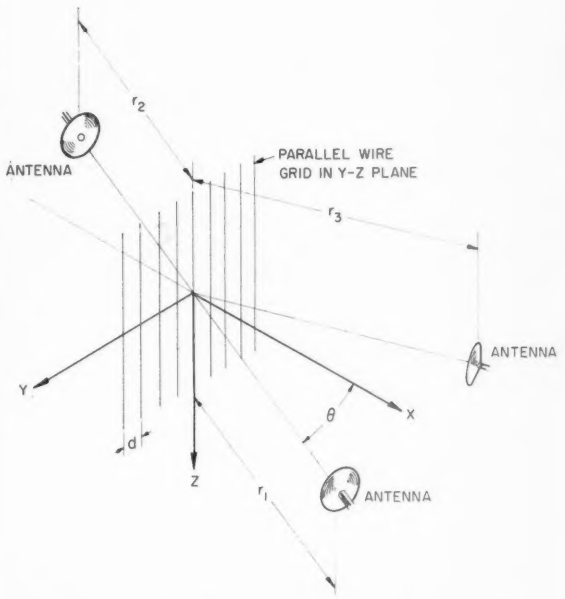


FIGURE 1. *Coordinate system for parallel wire grid.*

where

d=spacing between adjacent wires.
λ=wavelength.
α=radius of wires in the grid.

$F\left(\frac{d}{\lambda},\theta\right)$ is a correction factor which has been evaluated by MacFarlane.[2] A more general form is given by Wait.[1]

$\cos\theta Z_i\frac{d}{\eta_0}$ may be considered a correction factor involving the internal impedance Z_i of the wire. It is sufficiently small to be neglected in this case.

With the restrictions that $d>>\alpha$ and that $\lambda>2d$, R may be evaluated for the conditions of the experiment.

R^2 is then the proportion of energy reflected from the grid (reflectivity). Assuming no losses in the grid, the fraction of energy transmitted (transmissivity) is given by

$$T^2=1-R^2.$$

[1] J. R. Wait, Reflection at arbitrary incidence from a parallel wire grid, Appl. Sci. Research, B. **4**, 393 (1954–5).
[2] G. G. MacFarlane, Proc. Inst. Elec. Engrs. (London) Pt. IIIA **93**, 1523 (1946)

3. Grid Description

The parallel wire grid used in these experiments consisted of 75 copper wires (1.02 mm diam) spaced 1.5 cm on centers, making a total width of 111 cm. The wires were mounted on a frame approximately 3-m square as shown in figure 2. The frequencies used were 8.8 to 9.6 kMc, so that λ varies over the range of $2.08d$ to $2.27d$. In this case the ratio d/a is approximately 15. The wires were long enough so that shorting or opening their ends did not affect the transmitting or reflecting properties of the grid.

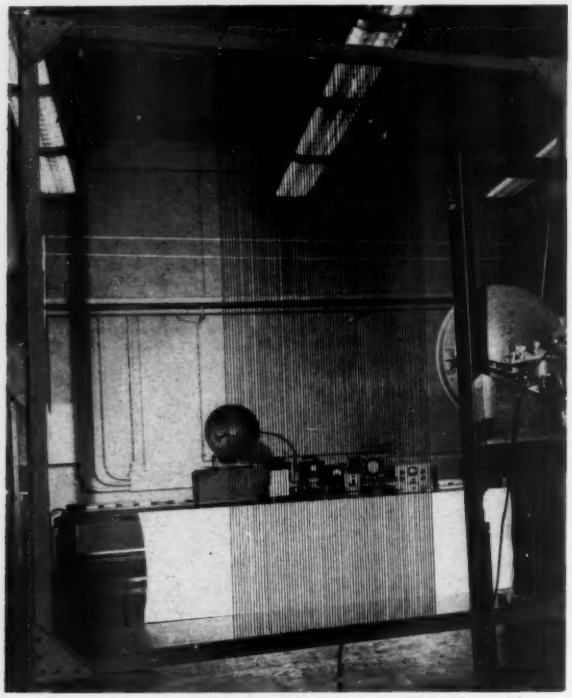


FIGURE 2. *Transmission through a parallel wire grid.*

4. Equipment

The measurements were made at X-band frequencies (8.8 to 9.6 kMc). A pulsed reflex klystron is used as a power source, with a ferrite isolator to minimize effects of the rest of the system on the power output and frequency of the source. The energy is fed through a frequency measuring device, precision attenuator and then through flexible waveguide to the transmitting antenna. Transmitting and receiving antennas are 18-in. parabolic reflectors. Slide-screw tuners are used at both antennas to provide a good impedance match, since reflections from the system terminals complicate the measurement process. These reflections are particularly bothersome when the transmitted wave is normally incident to an intervening obstacle, and techniques to eliminate their effect must be employed. Received energy is detected by the simple arrangement of a crystal in the waveguide at the receiving antenna. The detected signal (1,000 cps) is brought to an a-c amplifier and the output of this amplifier is rectified and fed to the recording meter.

The precision attenuator provides the standard of relative power measurement. The change in attenuation required to produce a given indication at the receiving terminal with and without an intervening obstacle is used to determine the power transmission coefficient of the obstacle. This procedure eliminates the need for assumptions of specific crystal characteristics at various operating levels. Power reflection coefficients are measured in a similar manner, comparing the power reflected from a surface to the power received over a direct transmission path of equivalent length.

The directional pattern of the main lobe of the 18-in. parabolic antennas was measured and the results are shown in figure 3. Here a relative power curve (in decibels) is plotted as a function of the H-plane angle from the antenna centerline, the apex of this angle being taken in the plane of the edge of the reflector. These measurements were made at a distance of 172 wavelengths (5.5 m).

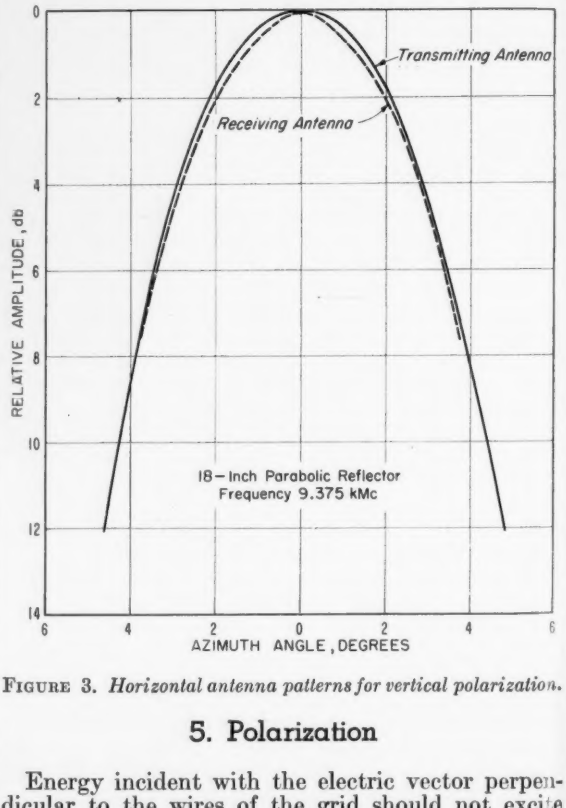


FIGURE 3. *Horizontal antenna patterns for vertical polarization.*

5. Polarization

Energy incident with the electric vector perpendicular to the wires of the grid should not excite currents of any appreciable magnitude in the wires and hence should not be reflected. This property was clearly demonstrated by experiment, the grid being completely transparent to horizontal polarization. This also indicates that the radiation from the antennas is linearly polarized.

6. Diffraction Effects

An indication of the grid size necessary to make edge effects negligible was obtained by moving the grid into the path as shown in figure 4, while observing simultaneously the transmitted energy. The initial rise in transmission above the free space value is due to diffraction as the edge of the grid approaches the path. When the edge of the grid has moved past the path centerline a distance of about 30 cm the transmission value becomes practically constant. The vertical bars at each of the measured points indicate the total range of variation of the data as the distance between transmitting and receiving antennas is varied from 4.5 to 7.5 m. A total minimum grid width of 60 cm would then be necessary for transmission measurement. This comment on grid size is applicable to an obstacle having attenuation comparable to the 1 to 2 db of the grid. If the attenuation in the obstacle were much higher the relative magnitude of diffracted energy could be greater.

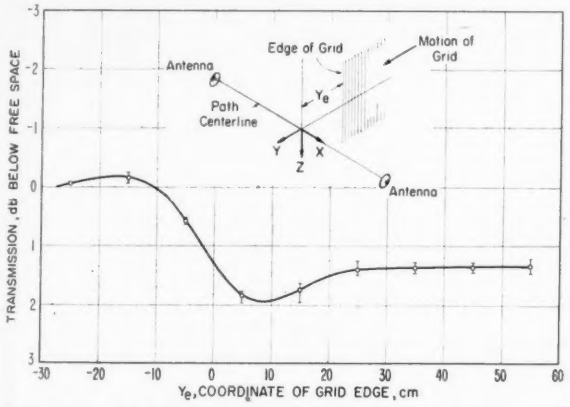


FIGURE 4. *Transmission through a parallel wire grid as grid is moved into path of antennas.*

7. Grid Transmission

A series of measurements of transmission through the 111-cm grid (1.5-cm spacing) were made at different frequencies and angles of incidence. As an example, figure 5 shows the energy transmitted through the grid (relative to transmission with no obstruction between the antennas) for a wavelength of 3.2 cm (9.375 kMc) and at angles of incidence up to 60°. The antennas were separated 5.5 m with the grid at the center of the path. The curve marked "calculated" was obtained by calculating the reflection coefficient as indicated previously, then assuming no loss in the grid. A series of these curves at different frequencies in the X-band is shown in figure 6. Here the data of figure 5 appear as the fourth set of curves at 9.375 kMc. The relation between transmission and angle of incidence is shown by each pair of curves (one measured and one calculated) while an indication of the transmission versus frequency at a given angle of incidence can be obtained by connecting corresponding angle-of-incidence points on each of the curves. The agreement of theory and

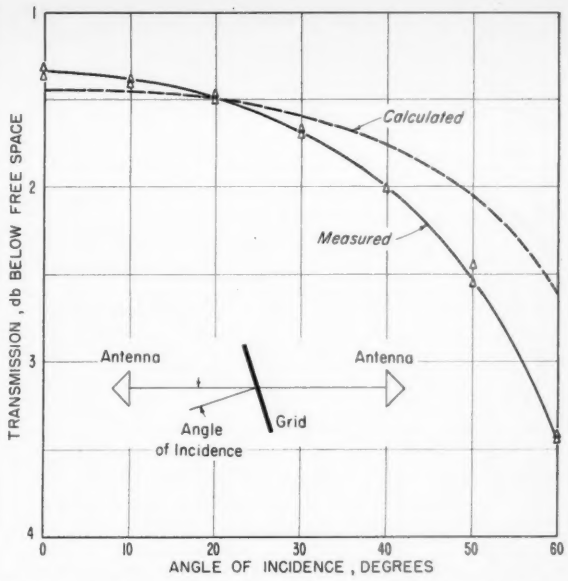


FIGURE 5. *Transmission through a parallel wire grid as a function of angle of incidence.*

measurement is fairly good at the lower angles of incidence, both with respect to angle of incidence and frequency. Discrepancy would be expected at the higher angles of incidence since at 60°, for example, the projection of the grid on a plane perpendicular to the path is 55 cm, corresponding to the 27-cm position of figure 4, where the projected grid size is becoming too small to agree with the infinite grid calculation. The lower transmission (higher reflection) also makes the diffracted energy relatively more important at these angles of incidence. It should also be noted that the calculation is made for a plane incident wave, while this condition is not completely met in the measurement.

8. Grid Reflection

Reflected and transmitted energy as a function of angle of incidence are shown in figure 7. The percent of reflected energy (reflectivity) was measured at a wavelength of 3.2 cm over a range of angles from 7° to 47°. Corresponding transmissivity data are shown for 0° to 60° incidence. The calculated curve may be considered to be either reflectivity or transmissivity on the assumption that there are no losses in the grid. The gap between the two measured curves is an indication of the portion of energy either lost in the grid itself or scattered in the directions other than the principal transmitted and reflected paths. This amounts to about 3 percent of the total energy for low angles of incidence, increasing to about 10 percent at the largest angle at which reflection measurements were made. This method does not separate the absorbed and scattered energy and hence is more effective in determining absorption from materials of fairly uniform composition and with plane surfaces.

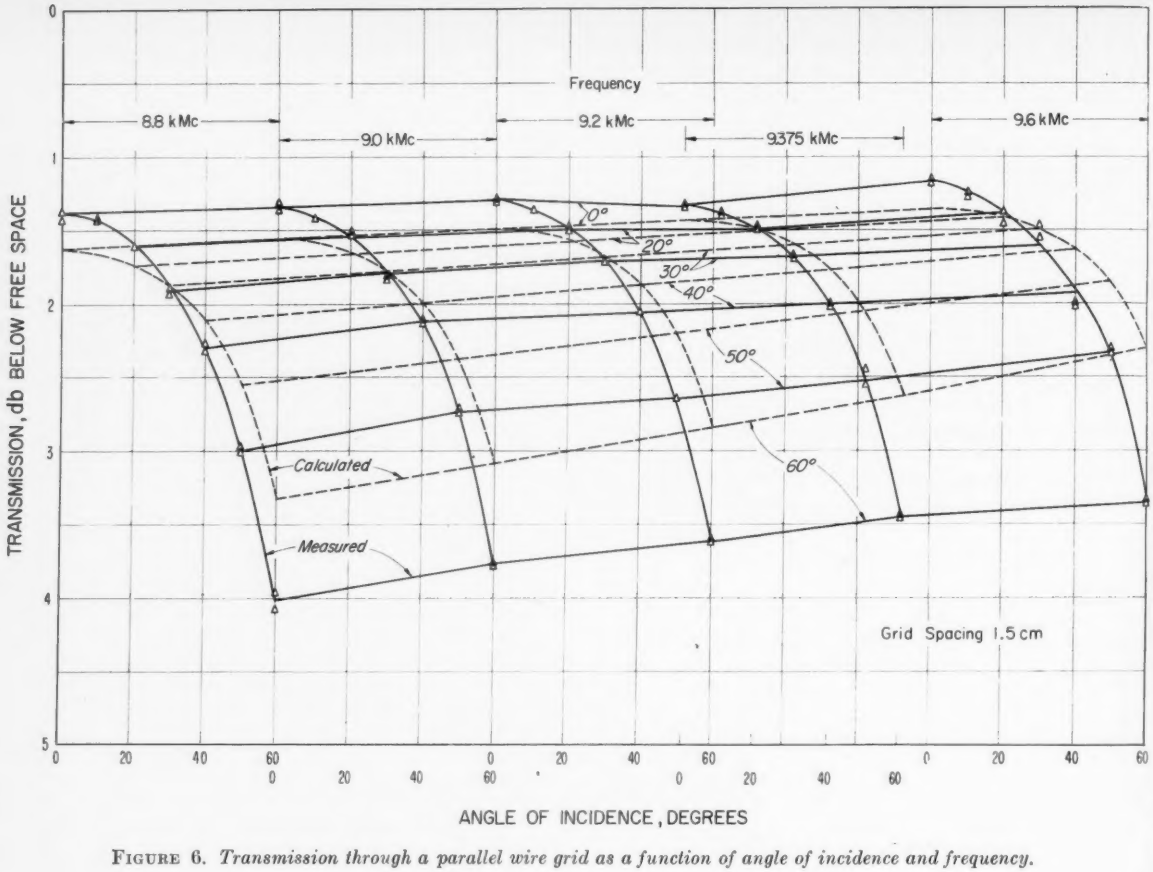


FIGURE 6. *Transmission through a parallel wire grid as a function of angle of incidence and frequency.*

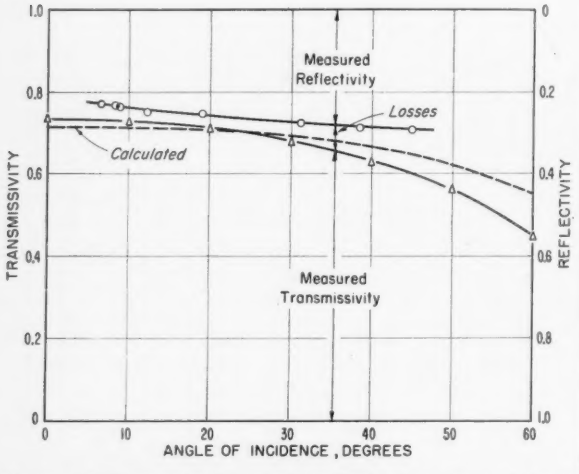


FIGURE 7. *Proportion of energy transmitted and reflected by a parallel wire grid.*

9. Conclusion

Transmission and reflection properties of a homogeneous material may often be determined from a knowledge of the dielectric constant and conductivity. If, however, the inhomogeneity of the material or the physical configuration of the material or radiating elements make calculation difficult, resort may be had to measurement. The agreement in theory and measurement obtained here indicates the degree to which the experimental configuration may be used in studying the properties of inhomogeneous structures. These methods are currently being applied over a wide frequency range in a study of various building materials and structures. The results may then be applied to cases in which it is impossible to avoid the presence of obstructing materials in an actual propagation path.

BOULDER, COLO. (Paper 63D1–10).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Synoptic Variation of the Radio Refractive Index*

B. R. Bean and L. P. Riggs

(January 22, 1959)

The synoptic variation of the atmospheric radio refractive index, evaluated from standard weather observations, is examined during an outbreak of polar continental air. It is found that the reduced-to-sea-level value of the refractive index is a more sensitive synoptic parameter than the station value. The reduced value is quite sensitive to the humidity and density structure of the storm under study while the great station elevation dependence of the station value tends to mask synoptic changes. The reduced value changes systematically with the approach and passage of the polar front. The present system shows a consistent increase of the reduced value in the warm sector of the wave and a marked decrease behind the cold front.

1. Introduction

The fact that the refractive index is readily determined from the standard weather observations taken in all parts of the world has led to many attempts to utilize it for the prediction of various properties of radio field strengths, particularly at frequencies in excess of about 30 Mc. It was noted early [1, 2] [1] that the seasonal cycle of vhf radio field strengths recorded far beyond the normal radio horizon was highly correlated with the refractive index. A further examination [3, 4] has shown that this correlation was about 0.9 when both variables were taken on a monthly mean basis but decreased rapidly for shorter mean periods. For example, 10-day means yielded a correlation coefficient of about 0.8, 5-day means about 0.7, and individual observations for a given hour each day throughout the year yielded a correlation coefficient of 0.5. The high correlation on a monthly mean basis is utilized by Norton [5] and Gray [6] to predict seasonal and geographic variations of radio field strengths. A recent experimental check of Norton's method with radio field strengths recorded in Japan [4] shows the same high correlation of about 0.9. The purpose of the present paper is to examine the synoptic variations of N_s and, although the correlations are not as high as those on a monthly mean basis, one may use Norton's or Gray's method as a useful guide of the day-to-day variation of radio field strengths.

The value of the refractive index of the atmosphere, n, may be determined from the expression for refractivity, N [7]:

$$N=(n-1)10^6=\frac{77.6}{T}\left(P+\frac{4810e_s RH}{T}\right), \qquad (1)$$

which has been shown to be in error by no more than 0.5 percent in the frequency range 0 to 30,000 Mc. Equation (1) may be evaluated from standard weather observations since it is a function of the atmospheric pressure, P, in millibars and the partial pressure of water vapor as expressed by the product of the relative humidity in percent, RH, and the saturation vapor pressure, e_s, for the temperature T in degrees Kelvin. When N is determined from the values of P, T and RH of the common surface weather observation it is denoted as N_s.

Another application of the instantaneous values of N_s is to the problem of radar elevation angle errors. It has been known for some time that the total bending or refraction of radio rays as they pass from the earth's surface upwards through the atmosphere may be expressed as a linear function of N_s [8] for rays departing at initial elevation angles in excess of about 10 deg. Recent statistical studies [9] have shown that this linear dependence of ray bending upon N_s may be extended to initial elevation angles of 0.5° with no regard to the refractive index profile, and to 0° for all N profiles but ducting. Further, recent papers [10, 11] have shown that a very effective model of the radio refractive index structure of the atmosphere for radio ray tracing purposes may be specified from a knowledge of the surface value alone. These latter applications are of direct utility when considering such problems as radar tracking of earth satellites or radio communications with space vehicles.

With the above applications as motivation plus a curiosity as to the synoptic potential of N, a rather detailed analysis of a single outbreak of polar continental air was studied.

The storm sequence chosen for analysis was a rather well-developed one with strong temperature and humidity contrasts between the air masses within the warm sector and behind the cold front. This strong contrast was one of the main reasons for studying this particular sequence, since, if the refractive index is not a good synoptic indicator in this case, then there is little hope of its being one in more subtle synoptic disturbances. The following sections will be concerned with examining the degree to which N does describe this particular storm.

Before proceeding with the N analysis of this storm, it might be well to review briefly the idealized distribution of air masses around a polar front wave. The region of interaction between the cold air of the poles and the warm air of the tropics is referred to as the polar front and is generally located between

*This work was sponsored in part by task 31 of the U.S. Navy Weather Research Facility, Norfolk, Va.

[1] Figures in brackets indicate the literature references at the end of this paper.

30° to 60° N. From time-to-time a section of the polar front is displaced northward by a flow of warm tropical air while an adjacent section is simultaneously displaced southward by a flow of polar air. The interaction of the flow of polar and tropical air results in the formation of a "wave" that moves along the polar front, often for thousands of miles. An example of a fully developed polar front wave is shown on figure 1(a) in the same manner that it would appear on the daily weather map. Across the great plains and eastern seaboard of the United States, the polar front wave normally moves along the line AB in figure 1(a). An idealized space cross section along the line AB is shown in figure 1(b). The warm tropical air that flows into the warm sector of the wave overrides the cool air before the wave to form the transition zone denoted as a warm front. The cold front represents the transition between the humid air of the warm sector that has been forced upwards and the advancing cold polar air. Squall lines are drawn to represent a belt of vigorous vertical convection and sharp wind shifts that frequently precede fast-moving cold fronts. The fronts and squall lines shown on the daily weather map represent the ground intersection of the transition zones between the various air masses.

The reader who wishes a critical appraisal of the present meteorological thinking on fronts, air masses, squall lines, etc., is referred to *Dynamic Meteorology and Weather Forecasting*, by Godske, Bergeron, Bjerknes, and Bundgaard, American Meteorological Society and Carnegie Institute of Washington, 1957.

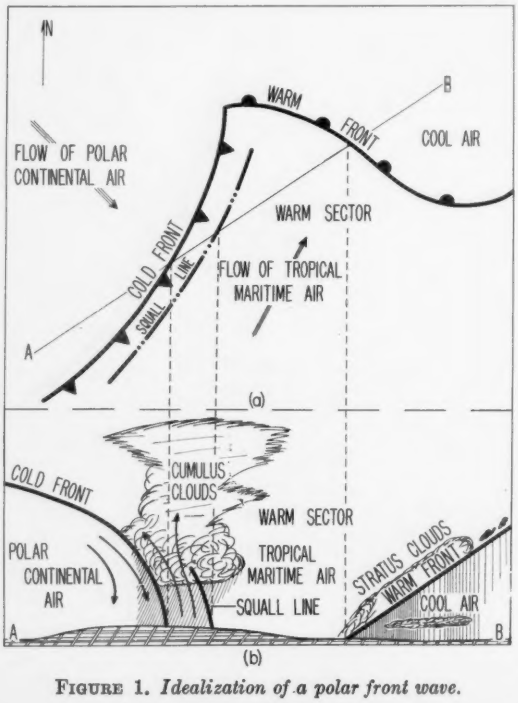


FIGURE 1. *Idealization of a polar front wave.*
Cross hatching represents area of rain.

2. Storm of 18 to 21 February 1952

The particular example chosen for study was the outbreak of polar air during the period 18 to 21 February 1952 during which a pronounced cold front developed and moved rapidly across the United States. The synoptic sequence is seen on figures 2 through 5 where contours of N_s are derived for four stages of this storm and compared to the superimposed Weather Bureau frontal analysis. This same procedure was followed throughout the present study; i.e., contours were derived first and then compared with the frontal pattern. The black dots on figure 8 indicate the 62 weather stations which were used to prepare the surface weather maps.

At the start of the sequence, a frontal system had come in from the Pacific Northwest and was composed of a cold front extending from northern Utah southwards into Arizona and a stationary front extending northeastwards into Wyoming. This frontal system was very slow moving for several days prior to the start of our sequence. With the outbreak of polar continental air east of the Rocky Mountains, this first front became more active and, as it moved ahead of the fast-moving polar continental front

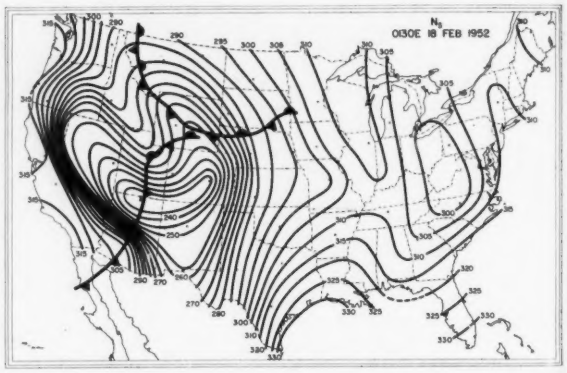


FIGURE 2. *Radio refractivity evaluated from Weather Bureau surface observations at 0130 e.s.t., 18 Feb. 1952.*

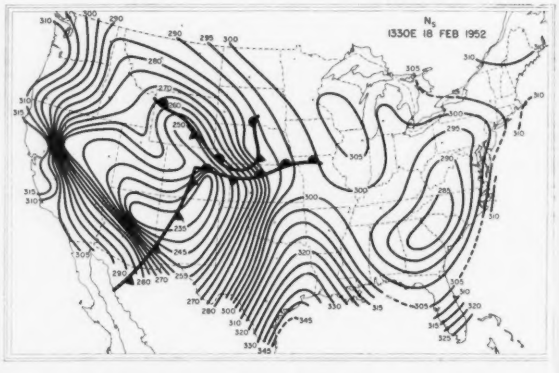


FIGURE 3. *Radio refractivity evaluated from Weather Bureau surface observations at 1330 e.s.t., 18 Feb. 1952.*

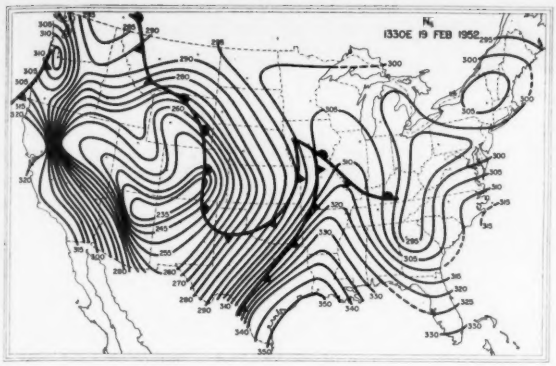


FIGURE 4. *Radio refractivity evaluated from Weather Bureau surface observations at 1330 e.s.t., 19 Feb. 1952.*

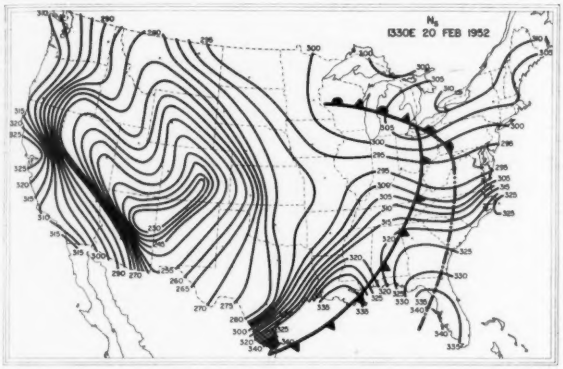


FIGURE 5. *Radio refractivity evaluated from Weather Bureau surface observations at 1330 e.s.t., 20 Feb. 1952.*

sweeping across the great plains, was reported as a squall line by the time it crossed the Mississippi River early on the morning of the 20th of February. During the latter part of the sequence, the polar maritime cold front—squall line was located in the developing warm sector of the polar front wave. The entire frontal system of cold front, polar front wave, and squall line then moved rapidly to the east coast by the morning of the 21st, thus completing the sequence.

3. *N* as a Synoptic Indicator

The dominant features of the N_s contours of figures 2 to 5 remain essentially unchanged by the passage of the frontal systems and the intrusion of cold air into the central part of the country. This is evident from the extremely time-stable map features such as the steep N_s gradient across the Sierra Nevada Mountains and the drop of N_s value with increasing elevation across the great plains between the Mississippi River and the Rocky Mountains. The semipermanent tongue of high values of N_s intruding up the Mississippi River valley does show an intensification and weakening with the passage of the frontal systems. The conclusion one reaches, then, is that N_s as such is not a particularly sensitive indicator of synoptic weather patterns.

This time stability of N_s patterns bears a rather marked similarity to contours of ground elevation above sea level, and thus one might suspect this height dependence to be associated with the pressure term in N. Meteorologists have long recognized that contours of station pressure tend to mask any synoptic pattern occurring over rapidly changing terrain. The variation of station pressure with altitude is so great compared with that associated with tropospheric storms that it is necessary to effectively remove this height dependence by reducing the station pressure to sea level in order to see more clearly the pressure variations due to the storm. A similar situation seems to exist with respect to N_s. To illustrate that this is indeed the case, let us restrict our attention to the data of 1330 e.s.t. on the 19th of February (fig. 4) where the polar front is located over the midwest. By separating N_s into a dry component, D_s, given by

$$D_s=\frac{77.6P}{T}, \tag{2}$$

and a wet component, W_s, given by

$$W_s=\frac{77.6(4810)e_sRH}{T^2}, \tag{3}$$

we will be able to discuss the height dependence more thoroughly since the station pressure (which is very height dependent) is contained in the dry term and the water-vapor pressure (which is sensitive to synoptic air mass changes) is confined to the wet term. Contours of D_s and W_s are given on figures 6 and 7 for 1330 e.s.t. on the 19th. Compare these contours with those of N_s for the same time (fig. 4). Note the similarity between N_s and D_s contours across the northern great plains and across the Sierra Nevada Mountains. It is evident that D_s and N_s have a similar dependence upon station elevation. Significantly, D_s contours do not seem to reflect the pressure pattern of the frontal system in the midwest. The contours of W_s, however, show a very flat gradient throughout the mountain regions but a relatively large gradient around the polar front wave. In fact, the W_s contours are extremely suggestive of the flow of warm, moist air into the warm sector of a classical textbook example of a polar-front wave.

It would seem that the variations of the dry term due to tropospheric storms would be more clearly indicated if its dependence upon station elevation were minimized. An arbitrary method of accomplishing this was adopted by noting that D_s is proportional to air density and then determining that the sea-level dry term, D_0, could be written:

$$D_0=D_s \exp\{0.1057h\}, \tag{4}$$

where h is in kilometers and the coefficient 0.1057

was empirically determined from the density-height distribution of the NACA dry standard atmosphere. Contours of D_0 were determined for the same synoptic situation as figure 4 and are shown on figure 8. It is immediately noted that the strong station elevation dependence of D_s is not in evidence in the D_0 chart. Further, the D_0 contours about the polar front wave are extremely suggestive of the air density distribution that one would expect; namely, simultaneous intrusion of dense air in the cold air mass behind the cold front accompanied by a northerly flow of warmer, less dense air from the Gulf of Mexico into the warm sector of the frontal system.

The status of the study at this point is that N_s and D_s do not appear to be particularly as sensitive to synoptic changes as are D_0 and W_s. We might proceed by analyzing our particular synoptic sequence in terms of D_0 and W_s but this seems unnecessarily complex. A simplification that suggests itself is to reduce N_s to sea level by the very same reduction factor as used on the dry term:

$$N_0=(D_s+W_s)\exp\{0.1057h\}. \qquad (5)$$

The contours of N_0 for the same data used on figures 4, 6, 7, and 8 are shown on figure 10. These contours of N_0 seem to exhibit the desired properties of a synoptic parameter; namely, a relative insensitivity to variations in station height combined with a marked sensitivity to the density and humidity structure of tropospheric storms. Thus the following analysis of the polar outbreak of 18 to 21 February, 1952 will be in terms of N_0. A further advantage of the use of N_0 is that N_s may easily be recovered from it for any practical application such as the determination of total atmospheric refraction.

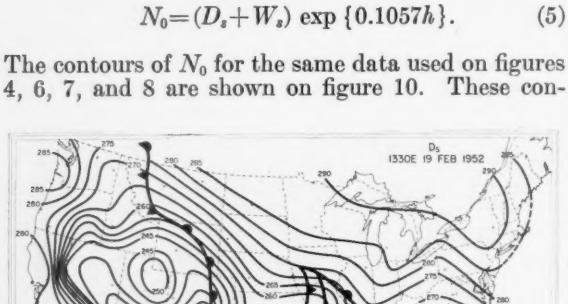


FIGURE 6. *Air density component of the radio refractivity evaluated from Weather Bureau surface observations at 1330 e.s.t., 19 Feb. 1952.*

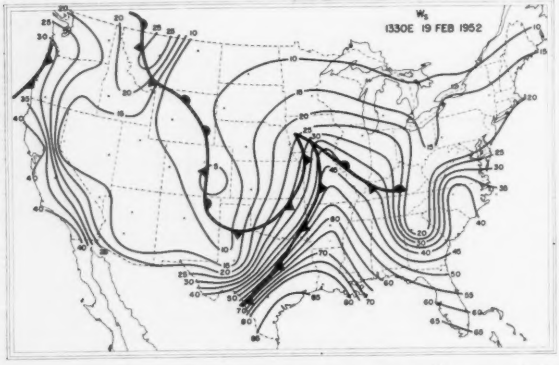


FIGURE 7. *Water vapor component of the radio refractivity evaluated from Weather Bureau surface observations at 1330 e.s.t., 19 Feb. 1952.*

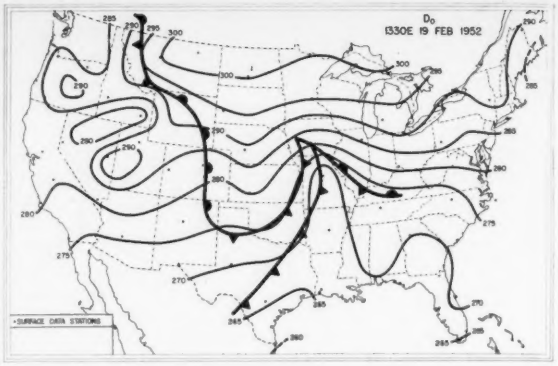


FIGURE 8. *Reduced-to-sea-level value of the air density component of the radio refractivity for 1330 e.s.t., 19 Feb. 1952.*

One might wonder at the advisability of arbitrarily reducing the wet term by the dry term correction factor. For the coastal areas of the country, where the exponential height correction factor is nearly one, this amounts simply to adding the D_0 and W_s maps while for the mountain areas, where the height correction is large but W_s is small, it amounts to simply reducing D_s to D_0 with the result that the gradiant of N isopleths obtained from the D_0 and W_s maps is essentially maintained on the N_0 maps.

A further advantage of N_0 over N_s is that it is obviously more simple to prepare charts of N_0. The contours of N_0 are more widely spaced and there are fewer of them.

4. N_0 Analysis of the Storm of 18 to 21 February 1952

Charts of N_0 were prepared from the Weather Bureau surface observations taken every 24 hr from 1330 e.s.t., February 18 until 1330 e.s.t., February 20, 1952, or, in other words, the period of time that it took the polar front wave to develop and move across the country. The N_0 contours are compared for these times on figures 9 to 11. These charts appear to indicate that the cold front that originally extended from Utah southward is a rather weak one since it does not seem to have much N_0 change associated with it. In the early stages of our sequence, this lack of air mass contrast is evidenced by the slight change of the position of the N_0=290 contour encircling west Texas and New Mexico as the frontal system moves through that area. By comparison, the cold front that sweeps down across the great

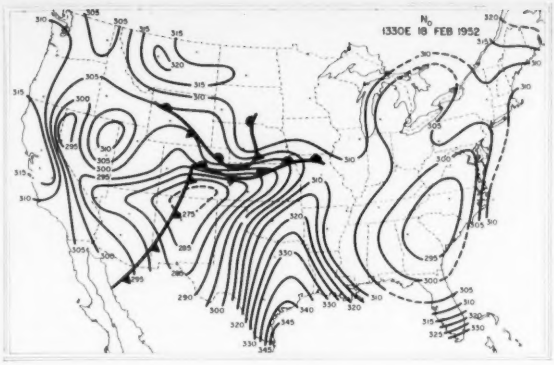


FIGURE 9. *Reduced-to-sea-level radio refractivity, 1330 e.s.t., 18 Feb. 1952.*

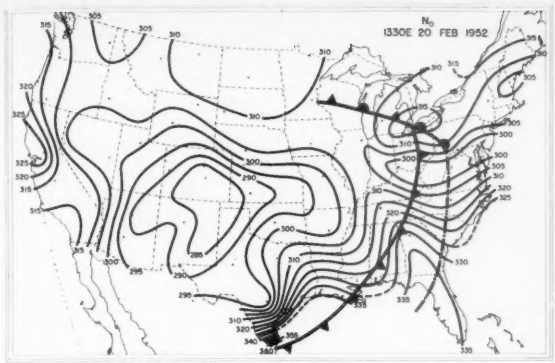


FIGURE 11. *Reduced-to-sea-level radio refractivity, 1330 e.s.t., 20 Feb. 1952.*

plains has a rather marked N_0 gradient across the front, due in large measure to the northward flow of warm moist air that forms a definite warm sector by 1330 e.s.t. on the 19th. It is perhaps significant that the N_0 contours indicate that the various frontal systems are transition zones rather than the sharply defined discontinuities of the text book examples, a point that has been recently enlarged upon by Palmer [12].

The variation of N_0 due to the frontal system can be seen on figures 12 and 13 where examples of the 24-hr changes of N_0 have been contoured. The 24-hr change, designated ΔN_0, was obtained by subtracting from the present value the value of N_0 observed 24 ago. The change was determined on a 24-hr basis in order to remove any effects of the diurnal cycle of N_0. The ΔN_0 charts show a rise of N_0 in the warm sector and a drop in N_0 behind the front amounting to $+35$ N units in the warm sector by 1330 e.s.t. on the 20th accompanied by a 40 to 50 N-unit drop behind the front. The relative sensitivity of N_0 to humidity changes is emphasized by the ΔN_0 charts.

Consider the N drop behind the cold front. This is a region of increasing pressure and decreasing temperature—a combination that increases the dry term while the rapidly dropping dewpoint decreases the wet term. The decrease in the wet term more than compensates for the increased dry term. As an example, in the 24-hr period ending 0130 e.s.t. on the 19th, the station pressure at Oklahoma City increased 13 millibars, the dry term increased 12 N-units while the wet term dropped 42-N units giving a net change of -30-N units. This N_0 rise in the warm sector and drop behind the cold front is consistent throughout the development of the polar front wave and appears to be what one would expect for this type of weather system.

The preceding maps appear to indicate that N_0 does reflect frontal system changes and that, with extensive studies of other cases, a system of objective forecasting of this parameter might be developed. Although this conclusion is based upon a study of a single storm, it appears to be borne out by preliminary analysis of other frontal systems.

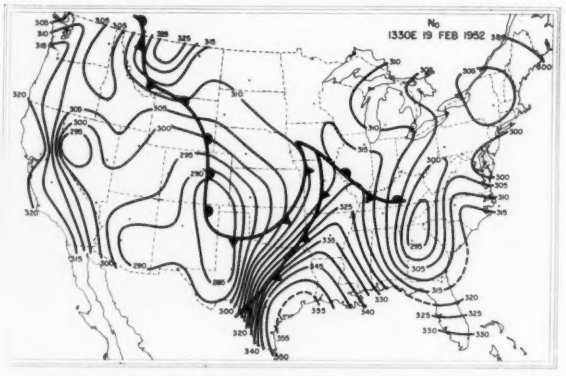


FIGURE 10. *Reduced-to-sea-level radio refractivity, 1330 e.s.t., 19 Feb. 1952.*

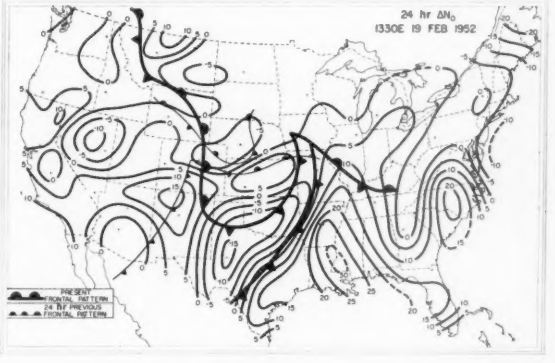


FIGURE 12. *Change in the reduced-to-sea-level refractivity for the 24-hr period ending 1330 e.s.t., 19 Feb. 1952.*

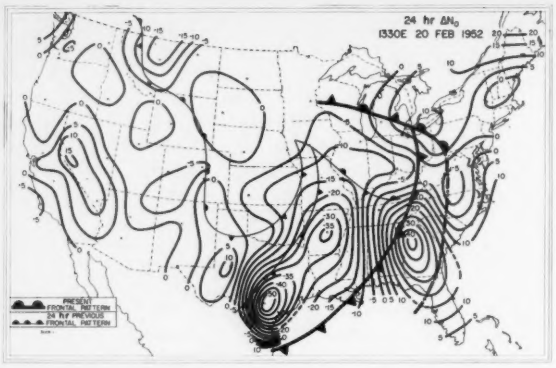


FIGURE 13. *Change in the reduced-to-sea-level refractivity for the 24-hr period ending 1330 e.s.t., 20 Feb. 1952.*

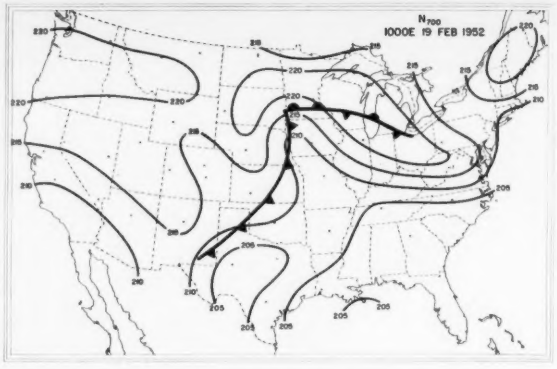


FIGURE 15. *Radio refractivity at the 700-mb level 1000 e.s.t., 19 Feb. 1952.*

5. Upper Air Analysis

Another facet of the synoptic variation of N is its three dimensional change with the passage of frontal systems. The 850-mb (5,000 ft above m.s.l.) and the 700-mb (10,000 ft above m.s.l.) charts were prepared for the times of radiosonde ascent (10 a.m. and 10 p.m. e.s.t.) throughout the synoptic sequence from the radiosonde reports of 43 U.S. sounding stations. It was not necessary to reduce the 850-mb or 700-mb level data since all of this data was already referenced to the indicated constant pressure level. Although as many maps of N at the 850-mb level, N_{850}, and 700-mb level, N_{700}, were prepared as there were N_0 maps, only two examples are chosen to illustrate the N changes aloft. The N_{850} and N_{700} contours are shown on figures 14 to 17 while their respective 24-hr changes, ΔN_{850} and ΔN_{700}, given on figures 18 to 19.

These N_{850} charts show that the northerly flow of warm humid air within the warm sector that was so prominent in the N_0 maps is also clearly in evidence at the 5,000-ft level. Further, a change pattern similar to that on the N_0 maps is also observed at the 850-mb level; i.e., a rise in N_0 value in the warm sector and a decrease behind the cold front. Surprisingly enough, by the time the frontal system is well developed, at 1000 e.s.t. on the 20th, the ΔN_{850} values are nearly as large as those on the surface.

The 700-millibar charts appear to be a bit more difficult to interpret than those of N_0 or N_{850}. It appears that at this altitude the wet term is usually negligible and N will normally vary inversely as temperature since the pressure is, of course, constant at the 700-mb level. By 1000 e.s.t. on the 19th an intrusion of low N values is observed into the 700-mb warm sector due to the advection of warm air northwards. The chart for 24 hr later is interesting in that it displays two N_{700}=225 highs, one between the squall line and the cold front and the other just south of the apex of the 700-mb wave. Interestingly enough, these two highs are due to quite different causes. The high centered over Atlanta appears to be due to the unusual transport of humidity to the 10,000-ft level since the 700-mb wet term at Atlanta increases from 4.5 to 25 N units in the 24-hr period ending with 1000 e.s.t. on the 20th. The second high centered over Omaha, appears to be due to an intense dome of cold air as indicated by the drop of the 700-mb temperature from −7.3° C to −21.4° C in the twelve hours preceding the map time. The ΔN_{700} charts show this change more clearly.

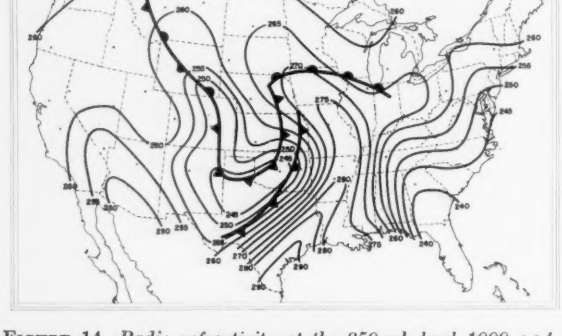


FIGURE 14. *Radio refractivity at the 850-mb level 1000 e.s.t., 19 Feb. 1952.*

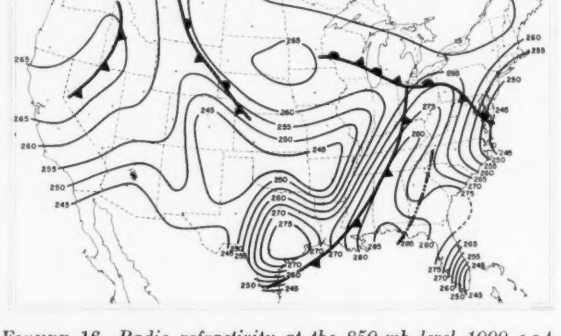


FIGURE 16. *Radio refractivity at the 850-mb level 1000 e.s.t. 20 Feb. 1952.*

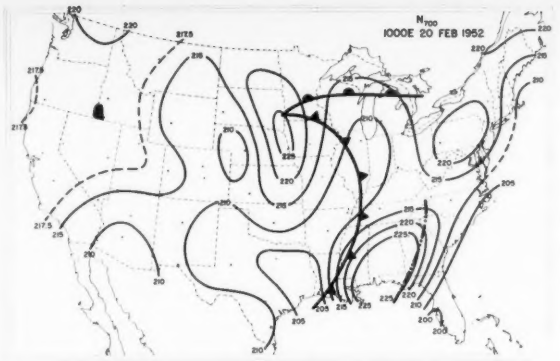


FIGURE 17. *Radio refractivity at the 700-mb level 1000 e.s.t., 20 Feb. 1952.*

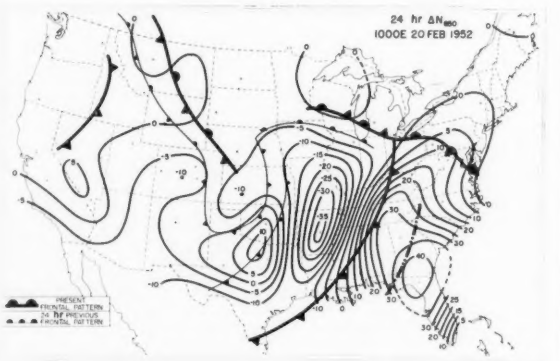


FIGURE 18. *Change in the radio refractivity at the 850-mb level for the 24-hr period ending 1000 e.s.t., 20 Feb. 1952.*

6. Conclusions

The present study indicates that the reduced-to-sea-level value of the refractive index is a better indicator of synoptic weather changes than the unreduced value.

A systematic increase of this reduced value is observed in the warm sector of a developing polar front wave accompanied by a decrease behind the cold front. The present system had about a $+35$ N unit rise and a 40 N-unit drop. This general pattern might be expected to occur in all fast moving cold fronts with varying intensity depending upon the individual synoptic pattern. In any case, it appears that the ΔN_0 pattern is a sufficiently stable and conservative property of the atmosphere so that it should be possible to develop forecasting rules for N_0 but not, of course, without analysis of many more N patterns.

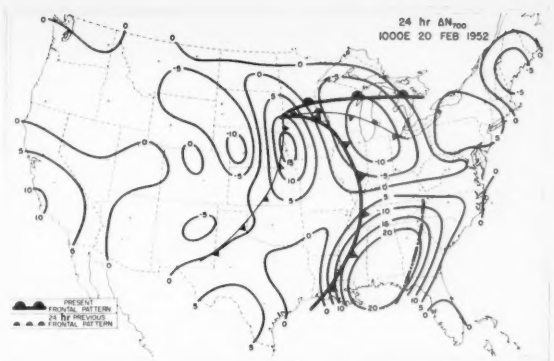


FIGURE 19. *Change in the radio refractivity at the 700-mb level for the 24-hr period ending 1000 e.s.t., 20 Feb. 1952.*

The authors express their appreciation to C. M. Miller for his preparation of the synoptic charts used in this report and to Miss G. Hoffmire for preparation of the manuscript.

7. References

[1] G. W. Pickard and H. T. Stetson, Comparison of tropospheric reception, J. Atmospheric and Terrest. Phys. **1,** 32 (1950).
[2] G. W. Pickard and H. T. Stetson, Comparison of tropospheric reception at 44.1 Mc with 92.1 Mc over the 167-mile path of Alpine, New Jersey, to Needham, Massachusetts, Proc. IRE *38* 1450 (1950).
[3] B. R. Bean, Some meteorological effects on scattered radio waves, IRE Trans. PGCS. **4** 32 (1956).
[4] M. Onoe, M. Hirai, S. Niwa, Results of experiment of long-distance overland propagation of ultra-short waves, J. Radio Research Labs. Japan **5** 79 (1958).
[5] K. A. Norton, Point-to-point radio relaying via the scatter mode of tropospheric propagation, IRE Trans. PGCS-**4,** 39 (1956).
[6] R. E. Gray, The refractive index of the atmosphere as a factor in tropospheric propagation far beyond the horizon, IRE Nat. Conv. Record pt 1, 3 (1957).
[7] E. K. Smith, Jr. and S. Weintraub, The constants in the equation for atmospheric refractive index at radio frequencies, Proc. IRE **41,** 1035 (1953).
[8] M. Schulkin, Average radio ray refraction in the lower atmosphere, Proc. IRE **40,** 554 (1952).
[9] B. R. Bean and B. A. Cahoon, The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles, Proc. IRE **45,** 1545 (1957).
[10] L. J. Anderson, Tropospheric bending of radio waves, Trans. Am. Geophys. Union **39,** 208 (1958).
[11] B. R. Bean and G. D. Thayer, On models of the atmospheric refractive index, in Proc. IRE **5,** 740 (1959).
[12] C. E. Palmer, Some kinematic aspects of frontal zones, J. Meteorol. **14,** 403 (1957).

BOULDER, COLO. (Paper 63D1–11).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 1, July–August 1959

Low-Frequency Propagation Paths in Arctic Areas

A. D. Watt, E. L. Maxwell, and E. H. Whelan

(December 19, 1958)

The very low ground conductivities encountered in arctic areas, and the particular ionospheric conditions prevailing at high latitudes, can lead to rather unusual radiation and propagation conditions. In order to determine the magnitude of these effects, field intensities from transmitters located in the Labrador and Greenland areas were measured both on the surface of the earth and during several aircraft flights over this area. The many factors involved in lf propagation are considered and calculated field intensities compared with experimental values. Under conditions where the initial portion of the propagation path is across icecap or permafrost, the attenuation observed is very great, and when the propagation path extends out over sea water, the field intensity recovery taking place after the coastline is crossed is very marked. Estimates of skywave field intensity appear to agree with the observed results provided the radiated field pattern is suitably modified by the antenna cutback factor which accounts for the presence of a finitely conducting curved earth. These vertical patterns based on work by Wait, along with the field intensity flight data, indicate that the siting of low-frequency stations several miles or more inland in arctic regions may cause a great increase in total transmission path loss.

1. Introduction

During July and August of 1957, field intensities were measured from lf transmitters located in the Labrador and Greenland areas with primary emphasis being placed on transmissions from Thule, Greenland at 98.5 kc and transmissions from Goose Bay, Labrador at 82.05 kc. The effective radiated power as a function of antenna current was obtained from each of the transmitters by observing the inverse field produced by the transmitter. All the field intensities presented are normalized to a transmitter radiated power of 28.7 kw. The actual field intensities on the surface were obtained by means of a 30-in. loop and a field intensity meter which was calibrated by a signal injected in series with the loop. The aircraft measurements were performed using the antenna on the aircraft themselves, which were in effect top loaded monopoles. The effective height in each case was obtained by calibrations on the ground before takeoff by comparing the voltages from the antenna with the field intensity obtained by means of a loop positioned approximately 100 ft from the aircraft.

2. Observed Field Intensities

The location of the transmitters and the surrounding terrain involved in the various flight paths are shown in figure 1. The field intensity observed for the groundwave along the Thule to Goose Bay path over the permafrost and icecap is shown in figure 2. The x's represent observations taken very close to the path, while the circles represent observations made at various other locations in the Thule area. It is interesting to note that even at distances less than 10 miles, the intervening terrain has an appreciable effect on the observed field intensity. In particular the two circles which lie on the inverse distance line at 5, 2, and 7.6 miles were for line-of-sight paths where there was a valley between the transmitter and the receiver. The next circle at 8.7 miles represents the field observed also on a line-of-sight path, but behind a glacier of about 1 mile extent in an open valley toward the transmitting site. It can be seen that this small icefield caused a 3-db decrease in field intensity. A similar abrupt decrease in field intensity was also observed when measurements were made on the edge and several miles out onto the main icecap. The two calculated groundwave curves were obtained from data by Wait [1][1] and it can be seen that the surface field apparently follows the groundwave curve for the poor arctic land up to a point where the glacier is reached. Beyond this it begins to approach that for ice in a manner as would be anticipated from the mixed path theory as has been described by Millington [2] and Wait [3, 4].

Figure 3 shows the field intensities observed during flight 1 and it is interesting to observe the rather smooth slow variations which occurred over land till the coastline was reached at Hopedale. From here until Saglek was reached at some 300 miles from the transmitter the field intensity varied quite rapidly as the aircraft flew over the rather rough coastline and it is believed that these variations were caused by coastline effects [4]. The very deep null in the vicinity of the Button Islands appears to have been caused by interference between the first hop skywave and the groundwave. The rather large recovery in field intensity out to the vicinity of Cape Dyer is apparently caused by the decrease in over land transmission caused by the path swinging out over the Davis Strait. From Cape Dyer to Cape Atholl, the field decreased in a rather smooth manner and

[1] Figures in brackets indicate the literature references at the end of this paper.

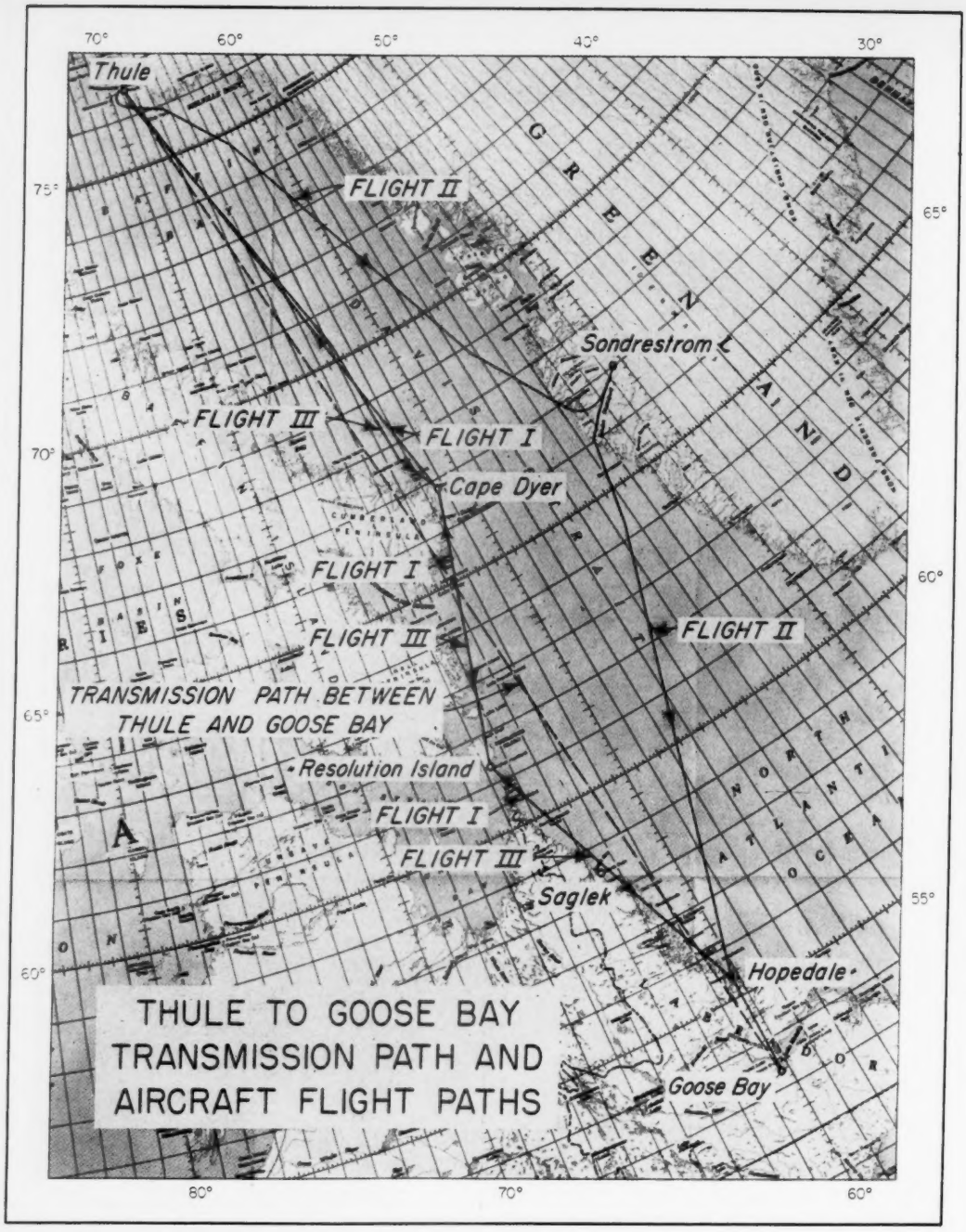


FIGURE 1. *Baffin Island area with flight paths and transmitting and receiving locations.*

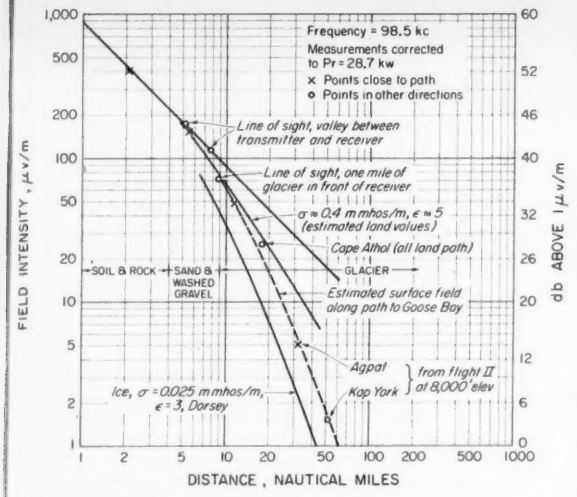

Frequency = 98.5 kc
Measurements corrected to Pr = 28.7 kw
× Points close to path
o Points in other directions
Line of sight, valley between transmitter and receiver
Line of sight, one mile of glacier in front of receiver
σ = 0.4 mmhos/m, ε = 5 (estimated land values)
Cape Atholl (all land path)
SOIL & ROCK
SAND & WASHED GRAVEL
GLACIER
Estimated surface field along path to Goose Bay
Agpat
Kap York
from flight II at 8,000' elev
Ice, σ = 0.025 mmhos/m, ε = 3, Dorsey
FIELD INTENSITY, μv/m
db ABOVE 1 μv/m
DISTANCE, NAUTICAL MILES

FIGURE 2. *Groundwave field intensities in the Thule area.*

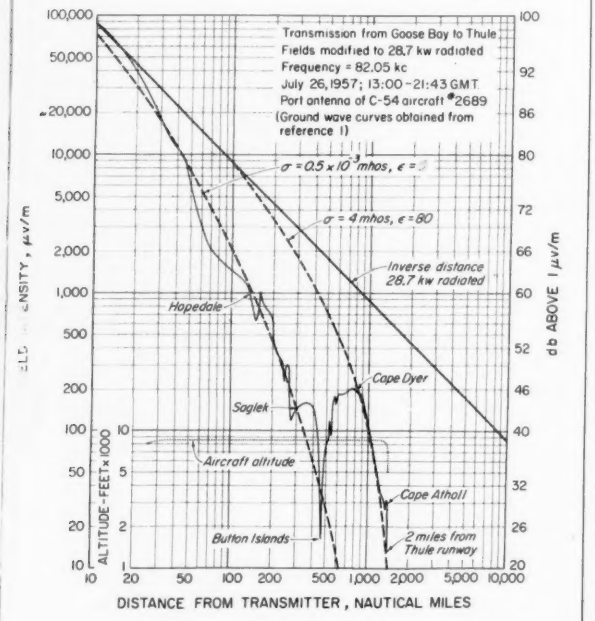

Transmission from Goose Bay to Thule
Fields modified to 28.7 kw radiated
Frequency = 82.05 kc
July 26, 1957; 13:00 - 21:43 GMT
Port antenna of C-54 aircraft #2689
(Ground wave curves obtained from reference 1)
σ = 0.5 x 10⁻³ mhos, ε = 3
σ = 4 mhos, ε = 80
Inverse distance 28.7 kw radiated
Hopedale
Cape Dyer
Saglek
Aircraft altitude
Cape Atholl
Button Islands
2 miles from Thule runway
ALTITUDE-FEET x 1000
db ABOVE 1 μv/m
DISTANCE FROM TRANSMITTER, NAUTICAL MILES

FIGURE 3. *Field intensities flight 1.*

no apparent effects were observed from flying over rather large patches of sea ice during this portion of the flight. Once Cape Atholl was reached, the field intensity began to vary at the coast and then decreased extremely rapidly until the last observation was made at 2 miles from Thule.

Figure 4 shows the field intensity from Thule as observed on flight 2. During the initial portion of this flight, the aircraft skirted along the coastline; however, once the aircraft departed from the coast, the field intensity rose considerably to a value

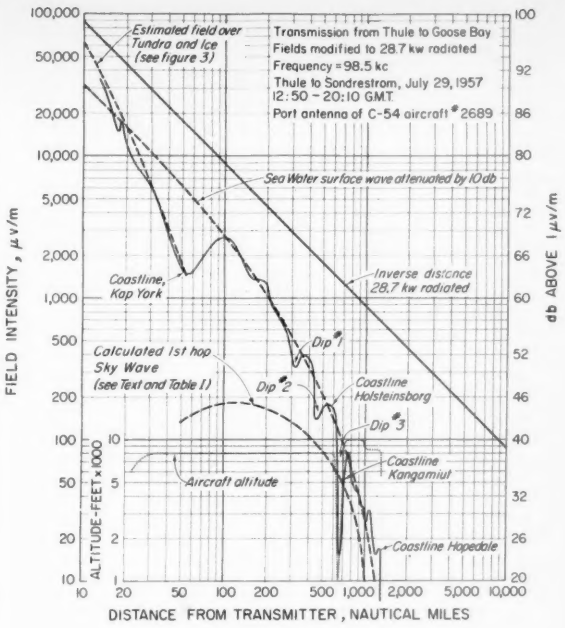

Estimated field over Tundra and Ice (see figure 3)
Transmission from Thule to Goose Bay
Fields modified to 28.7 kw radiated
Frequency = 98.5 kc
Thule to Sondrestrom, July 29, 1957
12:50 - 20:10 GMT
Port antenna of C-54 aircraft #2689
Sea Water surface wave attenuated by 10 db
Coastline, Kap York
Inverse distance 28.7 kw radiated
Calculated 1st hop Sky Wave (see Text and Table I)
Dip #1
Dip #2
Dip #3
Coastline Holsteinsborg
Coastline Kangamiut
Coastline Hopedale
Aircraft altitude
FIELD INTENSITY, μv/m
ALTITUDE-FEET x 1000
db ABOVE 1 μv/m
DISTANCE FROM TRANSMITTER, NAUTICAL MILES

FIGURE 4. *Field intensities flight 2.*

approximately 6 db higher than that observed at the coastline. Once this maximum recovery point had been reached, the field intensity is observed to drop in the manner anticipated for a surface wave over sea water from a transmitter with 10 db less power than that actually radiated. The dips at 320 and 450 nautical miles appeared to have been caused by interference with the first hop skywave, and although dip 3 may be partially caused by interference between sky and groundwave, it is likely that a large portion of the reduction is due to the poorly conducting mountainous terrain in the vicinity. In fact, it can be seen that the field intensity dropped very rapidly once the coastline was reached at Holsteinsborg, and that it remained low until the coastline was again reached at Kangamiut. Once the coastline at Hopedale was reached, the field intensity decayed very rapidly apparently due to propagation over the rather poor earth in this vicinity. This decrease was so rapid that the field intensity dropped into the noise level at 100 miles from Goose Bay.

Field intensities from the Sondrestrom transmitter at 132 kc were recorded during the second half of flight 2, and are shown in figure 5 where it can be seen that the field is apparently influenced to quite an extent by the surrounding terrain.

The results obtained on flight 3 (fig. 6) are interesting in that, although the flight was made during the day, the ionosphere was apparently disturbed to the point where it appeared as highly reflecting as at night. This can be seen from the general increase in field intensities observed during this flight compared to flight 2 and from the diurnal variations in

field intensity which were recorded at Goose Bay during, as well as preceding, this particular period. This particular flight also appears to contain a coastline dip with recovery effect; unfortunately, data was not taken that would record the exact depth of this dip. The rather high apparent ionospheric reflection coefficient present at this time raised the total field intensity in the region from 100 to 500 miles until a distance of 550 nautical miles was reached at which time there appears to be a canceling of energy received by the various modes. The lack of nulls up to this distance may be caused by a lower than normal layer height as well as by a possible fill-in from the 2d-hop skywave. After this null point was reached the field intensity recovered and remained rather high with some oscillations apparently due to coastline as well as interfering skywave effects until the coast was reached at Hopedale. At this point the field intensity again decreased very rapidly as in flight 2, and is probably due to the receiving antenna cutback effect, i.e., out-of-phase ground-reflected energy.

Since it was anticipated that the coastline effect upon the groundwave may be somewhat different than that on an incoming skywave, a flight was performed over the icecap starting at a distance of approximately 30 miles from the coast and extending several miles out over the sea. This flight was performed at 6,000 ft and the resulting field intensities from the Goose Bay transmitter, some 1,400 nautical miles distant, are recorded in figure 7. It can be seen that there is a very great change in amplitude in the vicinity of the coastline apparently due to interference which is probably caused partially by the discontinuity in conductivity as described by Wait [4], and in part by the change in

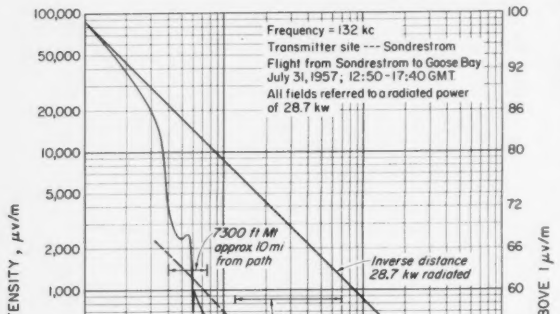

Frequency = 132 kc
Transmitter site --- Sondrestrom
Flight from Sondrestrom to Goose Bay
July 31, 1957; 12:50-17:40 GMT
All fields referred to a radiated power of 28.7 kw
7300 ft Mt approx 10 mi from path
Inverse distance 28.7 kw radiated
Mt in path 50 mi from transmitter
Coast
Sea water surface wave attenuated by 21 db
Hopedale
FIELD INTENSITY, μv/m
db ABOVE 1 μv/m
DISTANCE FROM TRANSMITTER, NAUTICAL MILES

FIGURE 5. *Field intensities flight 2B.*

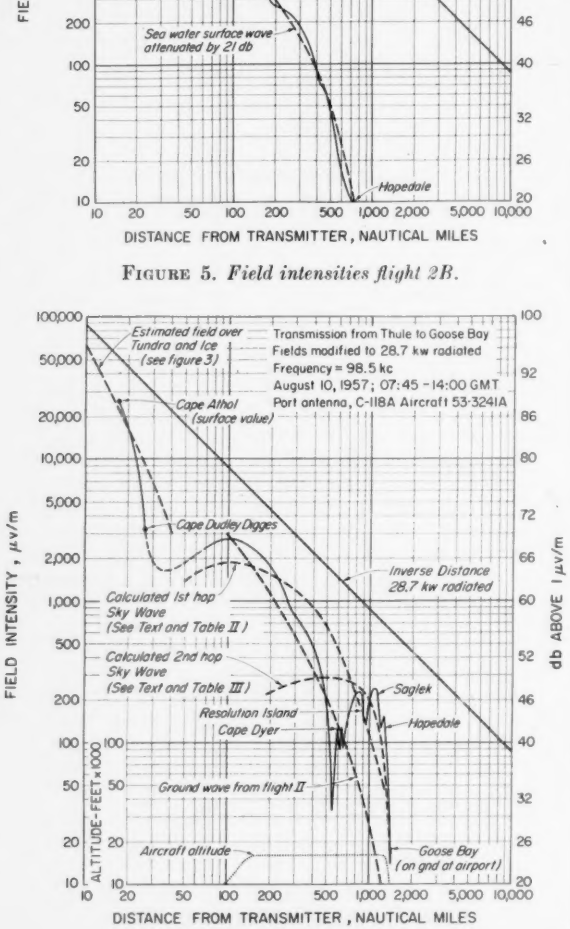

Estimated field over Tundra and Ice (see figure 3)
Transmission from Thule to Goose Bay
Fields modified to 28.7 kw radiated
Frequency = 98.5 kc
August 10, 1957; 07:45 - 14:00 GMT
Port antenna, C-118A Aircraft 53-3241A
Cape Athol (surface value)
Cape Dudley Digges
Inverse Distance 28.7 kw radiated
Calculated 1st hop Sky Wave (See Text and Table II)
Calculated 2nd hop Sky Wave (See Text and Table III)
Saglek
Resolution Island
Cape Dyer
Hopedale
Ground wave from flight II
Aircraft altitude
Goose Bay (on gnd at airport)
FIELD INTENSITY, μv/m
db ABOVE 1 μv/m
ALTITUDE-FEET×1000
DISTANCE FROM TRANSMITTER, NAUTICAL MILES

FIGURE 6. *Field intensities flight 3.*

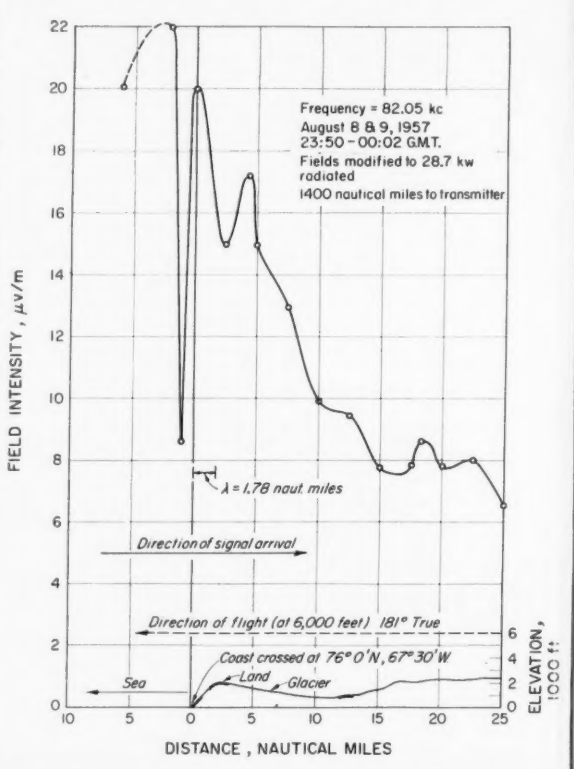

Frequency = 82.05 kc
August 8 & 9, 1957
23:50 - 00:02 GMT.
Fields modified to 28.7 kw radiated
1400 nautical miles to transmitter
λ = 1.78 naut miles
Direction of signal arrival
Direction of flight (at 6,000 feet) 181° True
Coast crossed at 76° 0'N, 67° 30'W
Sea
Land
Glacier
FIELD INTENSITY, μv/m
ELEVATION, 1000 ft
DISTANCE, NAUTICAL MILES

FIGURE 7. *Coastline effect on low-frequency carrier amplitude from a distant transmitter.*

elevation of the coastal cliffs as also described by Wait and Murphy [5]. The manner in which the field intensity decreases with increase in distance inland gives some idea of the loss in transmission or reception efficiency at this frequency for radio stations located inland in arctic regions. It should be emphasized that these observations were made at a height of 6,000 ft and that an even greater variation in field intensity is to be anticipated on the surface.

The manner in which the observed field intensities at Goose Bay vary with time can be seen in figure 8, where the fields were effectively averaged over approximately a 10 min period. The first 4 days of observation on the field from Thule appeared to have very similar diurnal patterns with a rather high maximum occurring approximately 1 hr after midnight. This rather short nighttime field is likely due to the fact that darkness exists for only a relatively short period of time at the 1-hop control point. During the day the field intensity decreased considerably and in general averaged around 2 μv/m.

A rather similar nightime maximum is observed on the 86-kc transmission from Keflavik although here the maximum peak occurred about 1½ hr before midnight as might be anticipated from the path location. Field intensity recordings on 98.5 kc were resumed on August 9 and the rather abrupt decrease just before midnight was observed as before; however, from this point on the pattern departed very markedly from that observed on previous days in that the field intensity remained rather steadily at a value almost 20 db above the normal daytime values. It was fortunate that we were able to obtain a complete flight from Thule to Goose Bay during the time when this abnormally high field intensity existed.

To our knowledge, records of solar flares during this period do not indicate a very marked disturbance on the 10th and it is rather difficult to explain these high field intensities. The fact that the high field intensities were from the Thule transmission and not noise or interference is rather well substantiated by independent measurements of flight 3 and the system performance figures during this period.

Solar records do show that a large flare occurred on the 11th followed by an extremely large ionospheric and magnetic disturbance on the 13th; however, this could explain the high field during the day on the 11th but cannot explain our observation on the 10th. It may be that abnormal ionospheric conditions of this type are very easily caused in auroral regions and occur frequently.

3. Analysis of Results

In attempting to explain the observed field intensities, we shall employ the results of numerous theoretical and experimental investigators. The method employed will be to start at the transmitting site and determine radiation and launching efficiency and then follow the energy as it is propagated to the point of observation.

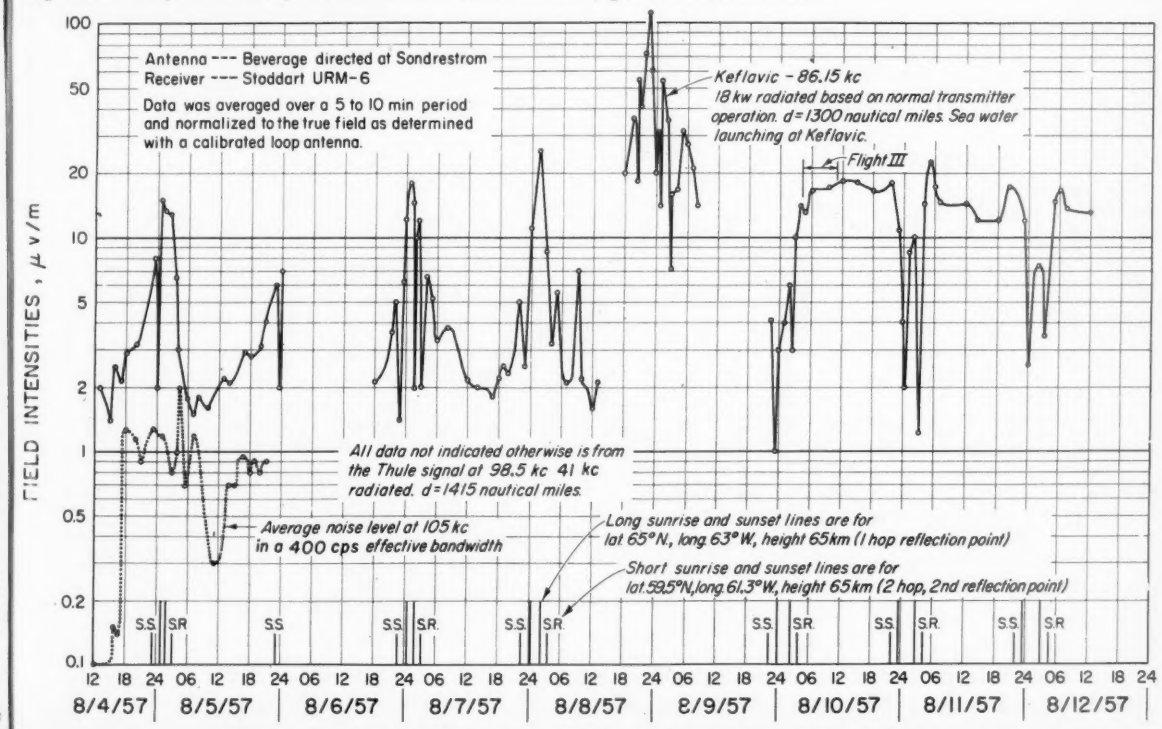

Antenna --- Beverage directed at Sondrestrom
Receiver --- Stoddart URM-6
Data was averaged over a 5 to 10 min period and normalized to the true field as determined with a calibrated loop antenna.
Keflavic - 86.15 kc
18 kw radiated based on normal transmitter operation. d = 1300 nautical miles. Sea water launching at Keflavic.
Flight III
All data not indicated otherwise is from the Thule signal at 98.5 kc 41 kc radiated. d = 1415 nautical miles.
Average noise level at 105 kc in a 400 cps effective bandwidth
Long sunrise and sunset lines are for lat 65°N, long 63°W, height 65km (1 hop reflection point)
Short sunrise and sunset lines are for lat 58.5°N, long 61.3°W, height 65 km (2 hop, 2nd reflection point)
FIELD INTENSITIES, μ v/m
100
50
20
10
5
2
1
0.5
0.2
0.1
S.S.
S.R.
12 18 24 06 12 18 24 06 12 18 24 06 12 18 24 06 12 18 24 06 12 18 24 06 12 18 24 06 12 18 24 06 12 18 24
8/4/57
8/5/57
8/6/57
8/7/57
8/8/57
8/9/57
8/10/57
8/11/57
8/12/57

FIGURE 8. *Diurnal field intensity variations as observed at Goose Bay, Labrador.*

3.1. Radiating Systems

The radiation efficiency for a given antenna is defined as the ratio of total power radiated to total power into the antenna. It is also well known that the radiation efficiency can be defined as follows:

$$\text{Rad eff} = R_r/R_t \tag{1}$$

where R_r is the radiation resistance and R_t is the total resistance seen at the base of the antenna. R_t can normally be measured while R_r can be calculated for relatively simple configurations [6, 7]. For a short vertical antenna we can write

$$R_r \approx 40^2\pi(h_a/\lambda)^2 \tag{2}$$

where h_a is the actual height of the vertical antenna and λ is the wavelength. If the form of the antenna is complicated by top loading or supporting masts, it frequently becomes convenient to express the radiation resistance as

$$R_r \approx 160\pi^2(h_e/\lambda)^2 \tag{3}$$

where h_e is now defined as the effective height of the antenna. Equations are available for calculating h_e for numerous types of antennas; however, it may become necessary to obtain this value by means of measurements in the field. This well-known method actually involves field intensity measurements to obtain the effective radiation power using the relationship

$$E = \frac{162.2\sqrt{P_r}}{d} \tag{4}$$

as described by Norton [8], which is good in the region d greater than λ and less than the point where the inverse distance relationship no longer holds. When the ground conductivity is unknown, it may be necessary to make observations at several distances to be sure that this distance is not being exceeded. If E is in millivolts per meter, d in nautical miles, the power radiated will be in kilowatts and it is relatively easy to then obtain the effective radiation resistance and the radiation efficiency of the antenna. It should be emphasized that eq (4) is based on the field from a short vertical antenna over a conducting surface, and that if the antenna height becomes greater than two tenths λ it will be necessary to include a correction factor because of the vertical radiation pattern. In the low-frequency range it is usually possible to reduce the antenna tuning coil losses and insulator losses to the point where they are rather small compared to the ground resistance. In Arctic regions where the ground conductivity is low this factor must be carefully considered. Methods of calculating the effective ground system resistance are described by Wait [7].

Normally it is possible to obtain sufficiently good ground systems so that with tall radiators the effecrive radiation efficiency is relatively high in the order of 50 to 90 percent.

3.2. Launching Efficiency

It is well known that the propagation of groundwaves over poor conductivity material causes a very rapid decrease in field intensity; however, the effect of poor conductivity materials on the launched skywave are not as well known. Since an appreciable amount of the terrain in these northern latitudes consists of ice, we have included some of its electrical properties in figure 9 as has been obtained from Dorsey [9], where the $-5°$ C curves are expected to approximate conditions on the Greenland icecap, and the $-50°$ C curves for those conditions on the Antarctic continent.

The usual procedure for calculating the skywave field launched from a given transmitter is to first obtain the free space field intensity as given by eq (5).

$$E_{fs} = \frac{81.1\sqrt{P_r}}{d} \tag{5}$$

where d is again in nautical miles, P_r in kw, E in millivolts per meter, and then modify it by the reflection coefficient for a plain earth similar to those contained in [5, sec. 10]. Recent work by Wait has provided us with antenna terrain cutback factors which are calculated for launching over a spherical earth, and include fields for negative launching angles such as occur over rather long propagation paths. These terrain cutback factors are presented in figures 10, 11, and 12 for various conditions. In figure 10 the effect of various types of terrain in the vicinity of the transmitting antenna are shown for 100 kc, and it is apparent that the presence of ice with its very low conductivity causes a tremendous decrease in the launching effectiveness of a transmitting facility. For example, if a skywave is being launched at an angle of 0 deg, which would correspond to a distance of approximately 1,200 nautical miles, it can be seen that the effective transmitter

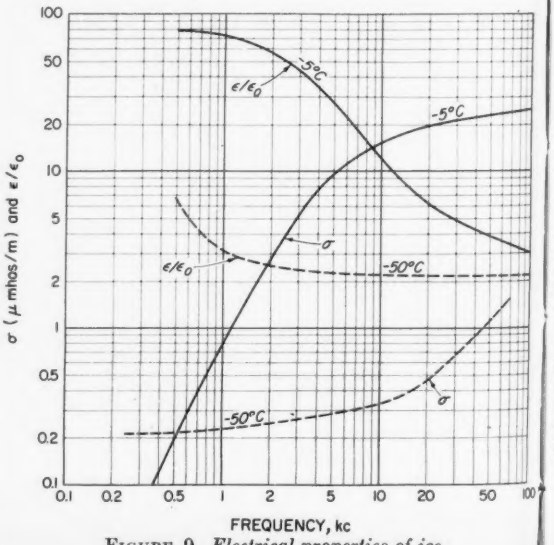


FIGURE 9. *Electrical properties of ice.*

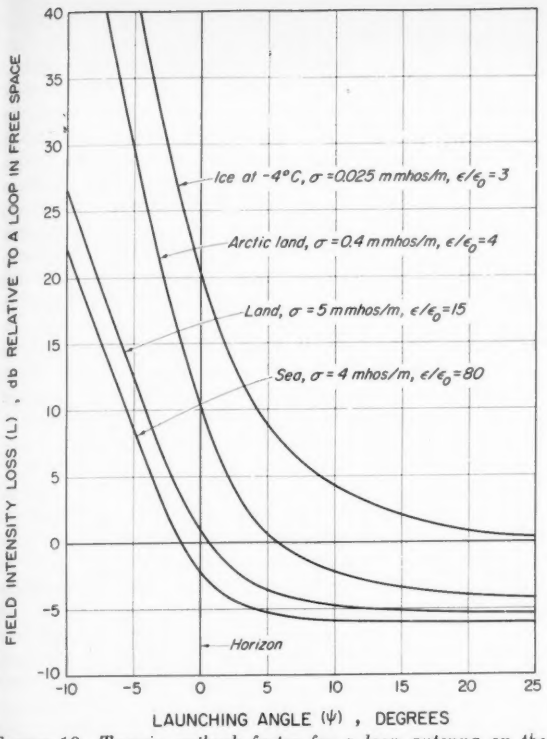
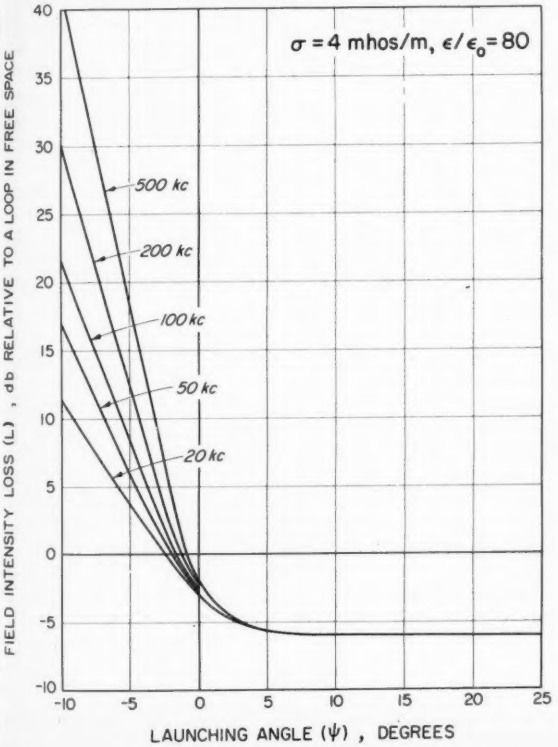

Ice at -4°C, σ =0.025 mmhos/m, ε/ε0 =3
Arctic land, σ =0.4 mmhos/m, ε/ε0 =4
Land, σ =5 mmhos/m, ε/ε0 =15
Sea, σ =4 mhos/m, ε/ε0 =80
Horizon
FIELD INTENSITY LOSS (L) , db RELATIVE TO A LOOP IN FREE SPACE
LAUNCHING ANGLE (ψ) , DEGREES

FIGURE 10. *Terrain cutback factor for a loop antenna on the surface at 100 kc (vertical E polarization) from Wait [9].*


σ =4 mhos/m, ε/ε0 =80
500 kc
200 kc
100 kc
50 kc
20 kc
FIELD INTENSITY LOSS (L) , db RELATIVE TO A LOOP IN FREE SPACE
LAUNCHING ANGLE (ψ) , DEGREES

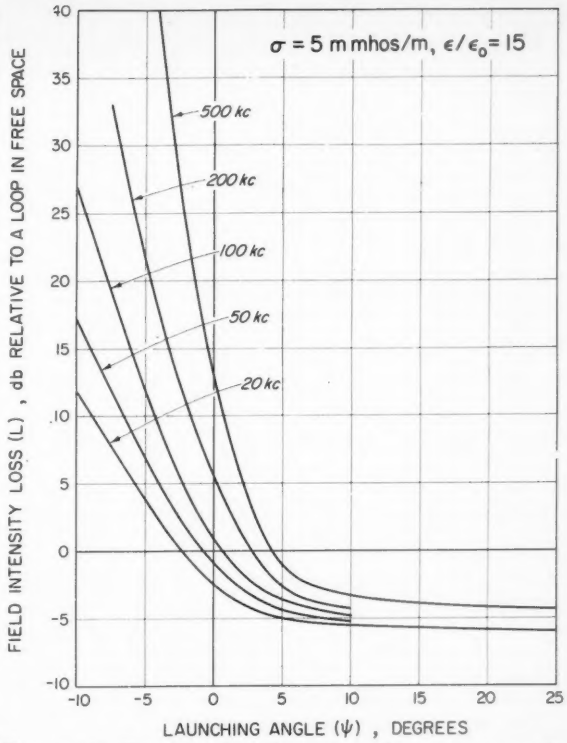

σ = 5 m mhos/m, ε/ε0 = 15
500 kc
200 kc
100 kc
50 kc
20 kc
FIELD INTENSITY LOSS (L) , db RELATIVE TO A LOOP IN FREE SPACE
LAUNCHING ANGLE (ψ) , DEGREES

FIGURE 12. *Terrain cutback factor for a loop antenna on the surface of average land (vertical E polarization) from Wait [9].*

power is reduced by a factor of 100 compared to launching over good land or sea water.

This launching factor must also be considered at the receiving location and the same type of correction made for the appropriate terrain in the vicinity of the receiver. Frequently the receiver may be elevated above the surface such as was the case of our aircraft measurements. In these conditions a height gain function as described by Norton [8] must be employed. Values of this function in the lf region for various types of terrain are shown in figure 13.

One of the important questions that is still not adequately answered concerns the amount of poor conductivity or good conductivity terrain in the vicinity of an antenna that is required to make the terrain cutback factor curves applicable. This is a rather complex problem, but some idea of the factors involved can be obtained from the analysis employed by Norton [11], in which he shows the method of calculating Fresnel reflection zones in front of an elevated antenna. When the incoming wave is an ionospheric reflected one approaching at very low angles and with an antenna on the surface, the calculations become rather difficult especially when a spherical earth is considered. When nearly grazing angles are considered at a frequency of approximately 100 kc, it would appear that distances in the

FIGURE 11. *Terrain cutback factor for a loop antenna on the surface of seawater (vertical E polarization) from Wait [9].*

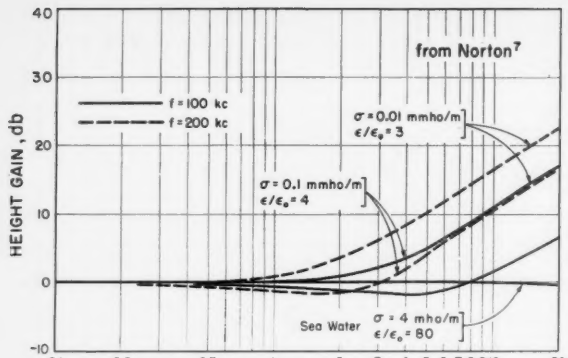


FIGURE 13. *Height gain function.* $G=\left|1+\frac{i2\pi h}{\lambda}\sqrt{\frac{i\epsilon_0\omega}{\sigma+i\epsilon\omega}}\right|$.

order of 30 miles may be important. An experimental investigation of this particular point can be seen with the aid of figure 7 where the amount of poor conductivity terrain in front of an elevated antenna was changed from 0 to 25 nautical miles. It is anticipated from other observations that the first hop skywave was essentially dominant at this time and that the incoming angle of arrival was approximately −1 deg. Since the icecap in this region is not extremely thick, it is possible that the appropriate terrain cutback factor should lie somewhere between the arctic land curve and that for ice. If this was the case, we would anticipate a change of approximately 18 db between the field observed over sea water and that inland as shown in figure 10. When we examine figure 7 we observe that the variation is closer to 11 db and we now must observe that this field intensity curve was taken at an elevation of 6,000 ft and that the height gain corrections must be applied. In figure 13 it is apparent that over sea water the height gain is essentially 0 db while over the rather poor arctic icecap a correction in the order of 6 db must be applied. When 6 db is subtracted from the 18 db difference in terrain cutback factors, we obtain a value of 12 db which is in much closer agreement with that observed. It was also observed that the largest amount of reduction in field intensity occurred in the first 10 to 12 miles or about 7 wavelengths from the coast. This type of reduction indicates, as is expected, that the terrain in the immediate foreground that is several wavelengths in front of the antenna, is much more important in determining the losses in launching efficiency than the ground further removed from the antenna itself.

3.3. Groundwave Propagation

When the media separating the transmitting and receiving points is homogenous, the surface wave field can be obtained very readily from references [1 and 7]. Millington [2] and Wait [3] have considered the mixed path problem and it can be seen that even in the simple case of only two conductivities along the path that the calculation becomes complex. When the distance separating the transmitter and the discontinuity in conductivity is relatively small, it is seen from figure 2 that the transition from the form of variation with distance characteristic of the first medium to that of the second is rather rapid. On the other hand, when the distance from the transmitter is large, as for example in flight 2, the recovery effect takes place over a much longer distance. This effect is very likely caused by the fact that close to the transmitter the field has a complex structure, while at longer distances the wave is somewhat characteristic of a plane wave. It should be mentioned that both the angle of propagation across the coastline and the height of observation will effect the rapidity of the recovery effect. After the transition period, from about 2 or 3 wavelengths to as much as 20 wavelengths, the groundwave is observed to behave in essentially the same manner as it would for a homogeneous medium with an increase or decrease in level depending upon the relative conductivities of the two media in question. As an example in figure 2 the field intensity appears to be following that anticipated for the over ice propagation at distances of 30 to 50 miles with an approximate 6-db increase in amplitude apparently caused by the 8 miles of soil in the vicinity of the transmitter itself. On the other hand, in figure 4 the surface wave increases from that at the coastline to a point where it follows the sea water surface wave attenuation curve that is attenuated by 10 db from that expected for a complete over water path. In other words, the 50 miles of tundra and ice in the vicinity of the transmitter have apparently caused a loss of 10 db for the surface wave. It would also be anticipated that had the terrain in the immediate vicinity of the transmitting antenna been glacier, that instead of being 10 db down, the field would have been approximately 16 db lower than for an all sea water path.

In figure 5 where transmission from Sondrestrom is observed at a frequency of 132 kc the surface wave is observed to decrease quite rapidly going through some oscillations in the vicinity of 50 miles apparently due to a 7,000-ft mountain off to one side. Once the coastline was reached, the course of flight was changed so as to bring the path back into the shadow of this mountain and the recovery effect observed was appreciably less than anticipated apparently due to the shading effect of the mountain. Here the over sea water surface wave appears to approach a decay rate anticipated for a surface wave 21 db below that for an all over sea water path. In this particular case the lossy media although not entirely glaciated, was very rough and produced an appreciable decrease in field.

A somewhat similar recovery effect was observed on flight 3; however, after the peak of the recovery was reached at approximately 100 miles from the transmitter, the rather high skywave reflection coefficient existing at this time caused an appreciable amount of skywave energy to be present so that it was impossible to determine from this flight the characteristics of the surface wave.

3.4. Skywave Field Calculations

The general ray path geometry involved in skywave propagation is shown in the sketch at the top of figure 14 where the launching and ionospheric reflecting angles are shown. These angles and the cosine of the ionospheric reflection angle are given in figure 14 for an effective height of 70 km which corresponds to typical daytime conditions and for an effective height of 90 km in figure 15 which is an effective nighttime condition. The dotted lines on the 1-hop psi curve show the anticipated values when atmospheric refraction is included in the calculations. This particular effect, along with a much more detailed description of the phenomena involved in skywave propagation, is described by Norton [12]. The time delays between the skywave and groundwave based on ray path length differences is given in figure 16 so that some idea of the separation between cancellation and reinforcement points can be calculated. This figure is also employed to determine the total length of the skywave which is used in calculating the inverse distance field.

It now becomes apparent that the skywave is reduced by an antenna vertical pattern and a launching factor at the transmitting site. It is also decreased by absorption and polarization conversion at the ionosphere and increased by a convergence coefficient as described by Norton [12], where the combined effect is considered as the reflection coefficient $_{||}R_{||}$ which can become greater than one for oblique paths. On arriving at the receiving point the skywave is changed by the launching factor and vertical antenna pattern at the receiver and the height gain factor if an elevated observation point is considered. For multiple hops the additional ionosphere reflection and conversion coefficients and ground reflections must be included. This can be

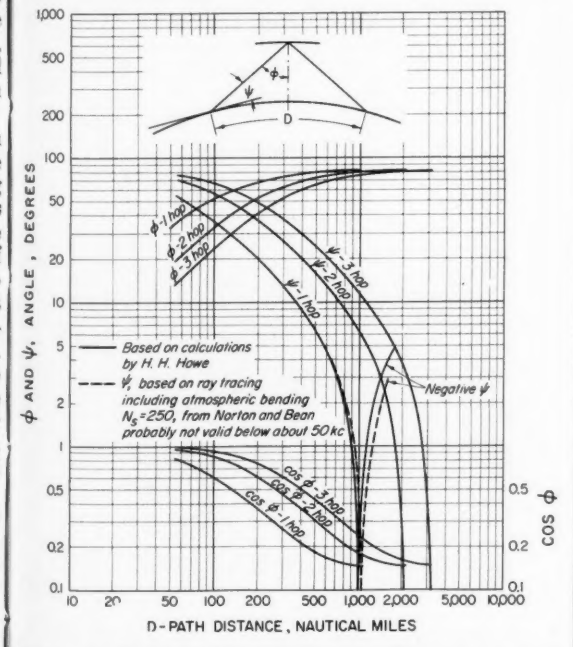

D
ϕ - 1 hop
ϕ - 2 hop
ϕ - 3 hop
ψ - 3 hop
ψ - 2 hop
ψ - 1 hop
Based on calculations by H. H. Howe
ψ, based on ray tracing including atmospheric bending N_s = 250, from Norton and Bean probably not valid below about 50 kc
Negative ψ
cos ϕ - 3 hop
cos ϕ - 2 hop
cos ϕ - 1 hop
ϕ AND ψ, ANGLE, DEGREES
COS ϕ
D - PATH DISTANCE, NAUTICAL MILES

FIGURE 14. *Ray path geometry and resulting angles for an effective layer height of 70 km.*

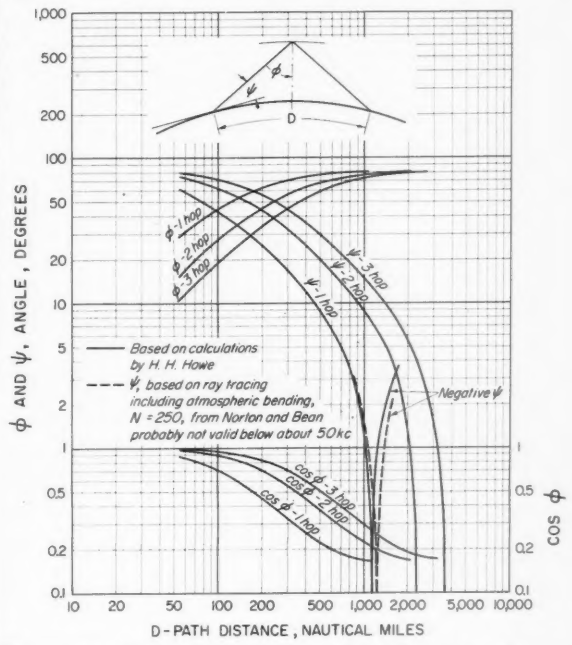

D
ϕ - 1 hop
ϕ - 2 hop
ϕ - 3 hop
ψ - 3 hop
ψ - 2 hop
ψ - 1 hop
Based on calculations by H. H. Howe
ψ, based on ray tracing including atmospheric bending, N = 250, from Norton and Bean probably not valid below about 50 kc
Negative ψ
cos ϕ - 3 hop
cos ϕ - 2 hop
cos ϕ - 1 hop
ϕ AND ψ, ANGLE, DEGREES
COS ϕ
D - PATH DISTANCE, NAUTICAL MILES

FIGURE 15. *Ray path geometry and resulting angles for an effective layer height of 90 km.*

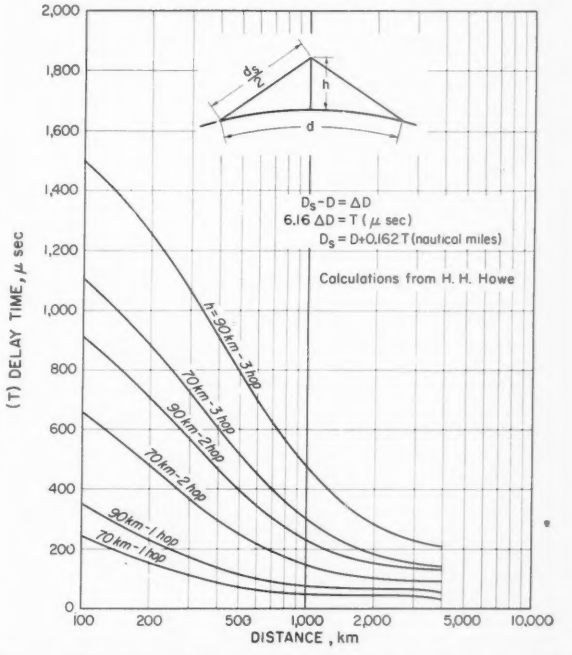

h
d
D_s - D = ΔD
6.16 ΔD = T (μ sec)
D_s = D + 0.162 T (nautical miles)
Calculations from H. H. Howe
h = 90 km - 3 hop
70 km - 3 hop
90 km - 2 hop
70 km - 2 hop
90 km - 1 hop
70 km - 1 hop
(T) DELAY TIME, μ sec
DISTANCE, km

FIGURE 16. *Transmission delay times between skywave and groundwave based on ray length only.*

summarized as shown in the following equation where instead of multiplying ratios, we shall add all the factors as expressed in decibels.

$$E_{sm}=K+P_r-D_p+A_t-L_t-I+C_m-(m-1)\ (I_m+G) -L_r+A_r+H \quad (6)$$

E_{sm}=mth hop skywave field intensity in decibels above 1 μv/m,
K=constant used in calculating free space field at unit distance,
=98.2 for nautical miles,
=99.4 for statute miles,
=103.6 for kilometers,
P_r=power radiated in db above 1 kw (see fig. 17),
D_p=inverse distance attenuation in db relative to unit distance (see fig. 18),
A_t=transmitting antenna free space gain in decibels at launching angle ψ (see fig. 19),
L_t=transmitting antenna launching loss in decibels relative to a loop in free space (see figs. 10, 11, 12),
I=ionospheric reflection loss=$-20\log_{11}R_{11}+C_m$ (see figs. 21, 22, 23),
m=number of hops,
C_m=convergence at ionosphere in db (see fig. 20),
I_m=ionospheric reflection loss for multiple hops which includes conversion as well as reflection coefficients,
G=ground reflection loss in decibels (see fig. 24),
L_r=receiving antenna launching loss in decibels relative to a loop in free space (see figs. 10, 11, 12),
A_r=receiving antenna free space gain at launching angle ψ in decibels (see fig. 19), and
H=receiving antenna height gain in decibels over surface value (see fig. 13).

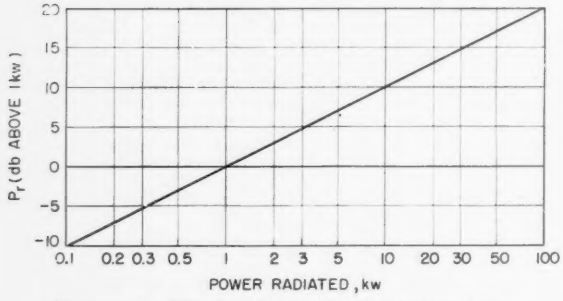


FIGURE 17. *Power radiated in decibels above 1 kw.*

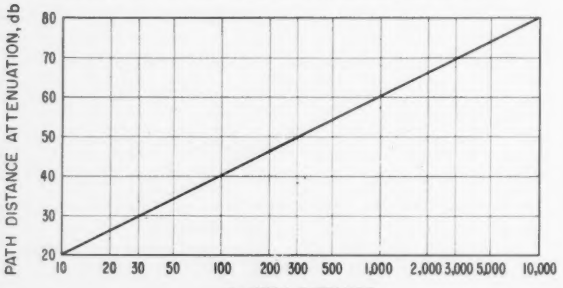


FIGURE 18. *Inverse distance.*

If the receiving antenna voltage is desired, it can readily be obtained by multiplying the field intensity in microvolts per meter by the effective height in meters.

Before attempting to apply this equation, we shall first consider the ionospheric reflection factor. It should be emphasized that the ionosphere is a very complex media, and that at present we only have an approximate knowledge of its structure. A recent survey of studies of the phase stability of ionospheric reflected radio waves from 16 to 500 kc has indicated

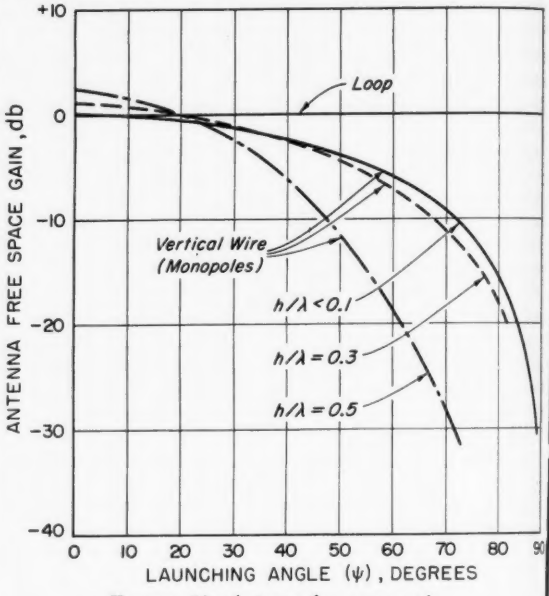


FIGURE 19. *Antenna free space gain.*

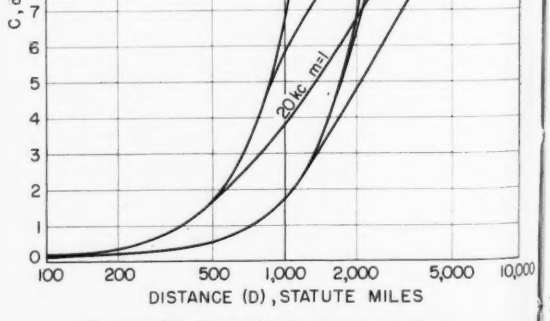


FIGURE 20. *Ionospheric convergence factor.*

h-70 km (maximum convergence anticipated under smooth ionospheric conditions).

that during the day the ionosphere for most of the time appears very smooth at least to 100 kc and probably beyond, and that at night the apparent roughness is only very slightly greater at vlf. At 100 kc and higher, the ionosphere may be smooth enough to yield essentially specular reflection and as a result full convergence; however, in the region of 500 kc the reflections may become fairly diffuse for an appreciable percentage (estimated 5 to 30) of the time.

On this basis, it can be seen that we should always employ the full value of C_m at vlf and also at lf during the day, and that it may also be applicable up to 500 kc for an appreciable amount of the time at night. Very excellent descriptions of many of the factors involved are given by Bracewell et al. [13] and Waynick [14]. Ionospheric reflection coefficients have been obtained by a rather large number of investigators [15 to 19] and their results along with some of our observations at Goose Bay and on flights 2 and 3 have been combined with appropriate corrections for antenna cutback factor where required and the results presented in figure 21. Here it is seen that the attenuation on reflection at 100 kc increases with the cosine of the angle of incidence to a maximum value at vertical incidence. It also shows that the summer noon values are much higher than winter noon which in turn are also higher than the nighttime values. It also appears that the variation in attenuation upon reflection as a function of angle is much less at night than during summer days.

To prevent us from reaching the erroneous conclusion that this behavior is typical at all frequencies, we have included figure 22 which shows the attenuation on reflection as a function of the cosine of the angle of incidence for 16.6 and 40 kc. Here it can be seen that there is an appreciable Brewster angle effect with a maximum attenuation in the region of 70 to 75 deg. This type of behavior is anticipated at vlf and the lower lf region below around 70 kc where the ionosphere effective boundary is apparently rather smooth and sharp as described in a theoretical paper by Wait and Perry [21]. As the frequency is increased, the boundary becomes less sharp so that the Brewster angle effect is not anticipated at frequencies of 100 kc and up. In order to determine the effect of different frequencies upon the attenuation on reflection, figure 23 has been prepared showing the loss for an angle whose cosine is equal to 0.15. This corresponds to a distance of approximately 800 nautical miles. It can be seen that the loss on reflection increases rather rapidly with frequency during summer daytime conditions reaching a value in excess of 70 db at a frequency of 700 kc. At noon in winter, the increase is less pronounced and during nighttime conditions the loss increases very little with frequency until 100 kc is reached and beyond this point a maximum value of 10 to 20 db in the region of 400 kc is expected. One of the interesting apparent indications of the data making up this particular curve are that at very low frequencies, in the order of 15 to 20 kc, the reflection

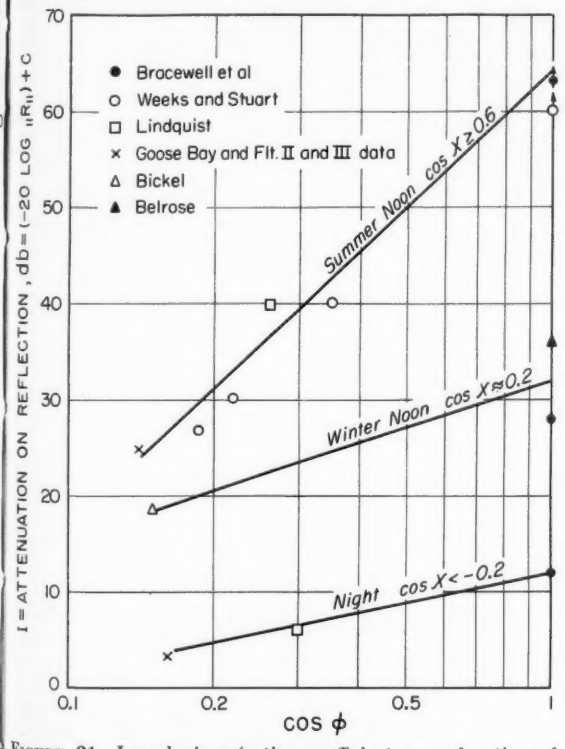

I = ATTENUATION ON REFLECTION, db = (−20 LOG ||(R||) + C
Bracewell et al
Weeks and Stuart
Lindquist
Goose Bay and Flt. II and III data
Bickel
Belrose
Summer Noon cos X ≥ 0.6
Winter Noon cos X ≈ 0.2
Night cos X < −0.2
COS φ

FIGURE 21. *Inospheric reflection coefficient as a function of angle of incidence at 100 kc*

Convergence factor removed. Estimated median values. Actual value at any time may differ appreciably from that shown.

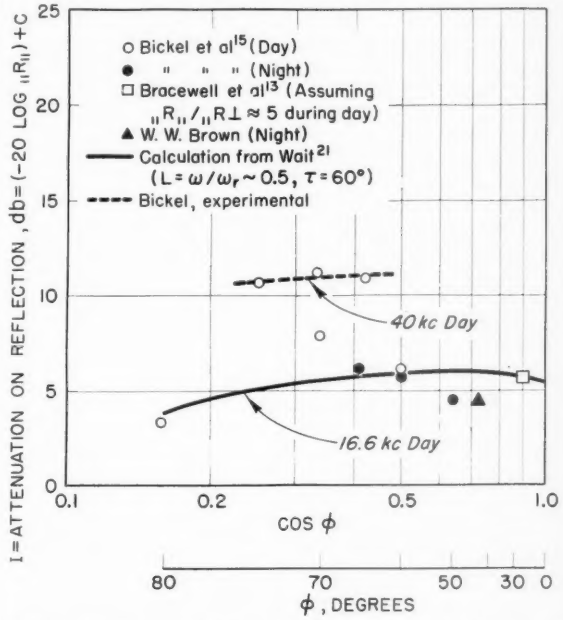

I = ATTENUATION ON REFLECTION, db = (−20 LOG ||(R||) + C
Bickel et al[15] (Day)
" " " (Night)
Bracewell et al[13] (Assuming ||R|| / ||R⊥ ≈ 5 during day)
W. W. Brown (Night)
Calculation from Wait[21] (L = ω/ω_r ~ 0.5, τ = 60°)
Bickel, experimental
40 kc Day
16.6 kc Day
COS φ
φ, DEGREES

FIGURE 22. *Ionospheric reflection coefficient as a function of angle of incidence at 16 and 40 kc.*

Convergence factor removed.

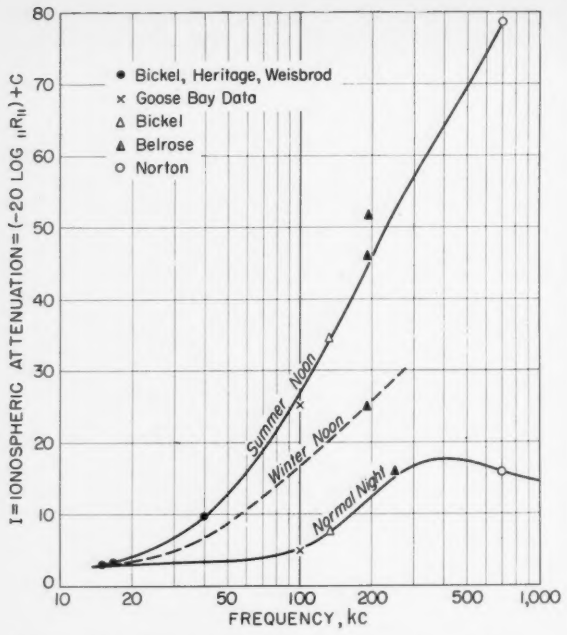


FIGURE 23. *Approximate ionospheric reflection coefficient at grazing (cos ϕ-0.15), (d-800 nautical miles), for vlf and lf.*
Estimated median values.

coefficient at near grazing angles does not vary greatly either with season or time of day. This at first might seem to contradict the amount of observed increase in attenuation during the day over that observed at night. A closer examination of the problem by Wait [21], reveals that in general the increase in daytime attenuation can be explained by the lower height of the ionosphere with the corresponding changes in angles of launching.

The seasonal and diurnal variation of the ionospheric reflection coefficient are well known to be correlated with solar radiation as has been shown for example by Pierce [23] where he relates the path attenuation coefficient to the sun's zenith angle. The approximate zenith angles corresponding to the conditions described in figures 21 and 23 have been included to give a rough idea of this effect. It

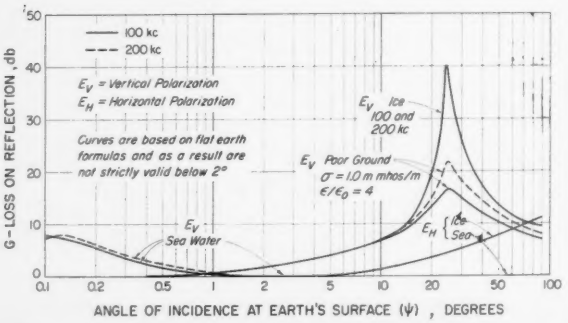

G - LOSS ON REFLECTION, db
ANGLE OF INCIDENCE AT EARTH'S SURFACE (ψ), DEGREES
100 kc
200 kc
E_V = Vertical Polarization
E_H = Horizontal Polarization
Curves are based on flat earth formulas and as a result are not strictly valid below 2°
E_V Ice 100 and 200 kc
E_V Poor Ground σ = 1.0 m mhos/m ε/ε$_0$ = 4
E_V Sea Water
E_H Ice
Sea

FIGURE 24. *Ground-reflection attenuation. See Terman [5].*

should be mentioned that sunspot activity will also affect the reflection coefficient and when more detailed curves of loss on reflection are prepared, they will undoubtedly include these effects. When more than one hop is considered, the ground reflection coefficient for each polarization must be considered as shown in figure 24.

When determining the efficiency of a given carrier in the presence of noise and interference, it is necessary to know the characteristics of the fading as well as that of the interference present as described by Watt et al [24]. Unfortunately, statistical distributions of the instantaneous carrier amplitude were not obtained during these observations; however, from the recorder record with a 1-cps band, the fade rate was observed to vary from fairly rapid excursions during the sunrise and sunset periods to slower variations during the day and short night periods. It was interesting to observe that the abnormally high field intensity periods recorded during the day of the 10th and 11th showed a very steady carrier level with essentially no fading at all. This would tend to indicate that the ionospheric reflecting layer present at this time was much steadier than normally observed.

Some of the fading characteristics as described by Bowhill [20] are shown in figure 25 where it is interesting to observe that the fade rate is generally very low at frequencies below 70 kc and that it increases at higher frequencies. Bowhill also points out that in the intermediate region of 60 to 100 kc that fading records show the presence of both a slow component of the magnitude of 0.01 cps and a high component whose rate is in the order of 0.2 cps. These fade rate observations combined with spaced receiver measurements indicate layer drifts in the ionosphere and rather random blob sizes which during the nighttime at a height of approximately 90 km appear to have an average effective horizontal size of 6 km. At a 100-km height as ob-

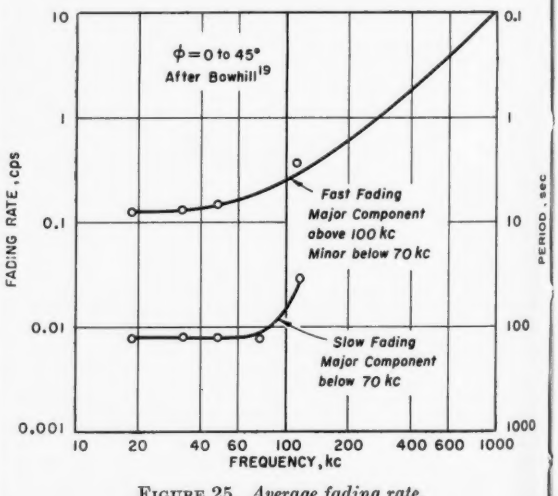


FIGURE 25. *Average fading rate.*
Nighttime observations at near vertical incidence.

served by frequencies in the order of 150 kc the average blob size appears to be 0.5 km. Somewhat similar observations have been made by Landmark [25] although the size of the blobs he reports are somewhat larger than those expected by Bowhill.

In any event it is interesting to note the change in appearance of the ionosphere as frequency is increased in the region of 70 kc and compare it with the nighttime ionospheric reflection coefficient presented in figure 23. Here at more nearly grazing incidence the effective change might be anticipated to occur at a lower frequency and it is possible that were sufficient data available, we would find that the reflection coefficient at night changed very little up to 30 or 50 kc and that above this region the attenuation would increase rather rapidly.

It should be pointed out that the factor affecting the amount of convergence obtained is actually the effective roughness rather than the horizontal blob size. Norton [12] has summarized the phase sta-

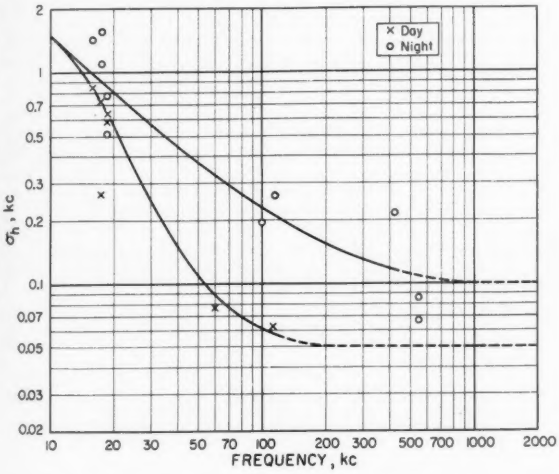


FIGURE 26. *Median effective ionospheric roughness parameter, σ_h, obtained from observations of σ_Ω.*

σ_Ω(degrees)$\equiv 2.4 f_{kc} \cos \phi \sigma_h$ (kilometers).

bility observations of a number of investigators over the frequency range of 16 to 500 kc and the results of this study plotted in terms of ionospheric roughness is shown in figure 26. These results indicate that the convergence shown in figure 20 is probably always valid up to 100 kc and valid during the day up to 500 kc. At night from 200 to 500 kc the amount of convergence present is likely to vary from zero to the full amount.

Three sample calculations for a 100 kc-frequency are included in this report and are shown in tables 1, 2, and 3. The results are also shown on figures 4 and 6. The calculations were made according to the existing conditions at the time flights 2 and 3 were made. For instance, flight 3 was made during daylight hours in northern latitudes where cos $\chi \approx 0.2$. Our measurements indicated however, that the ionosphere was not typical of daytime conditions,

TABLE 1. *Skywave field intensity calculation—Thule to Goose Path*

$E_{s1}=K+P_r-D_p+A_t-L_t-I+C_1-L_r+A_r+H$

Condition: Daytime, northern latitudes, cos $\chi \approx 0.2$, $I=I$ med+2 db

Distance nautical miles	50	70	100	200	400	700	1,000	1,250
ψ_1 (70 km)	57°	47°	37°	20°	9°	3.2°	0.3°	−0.5°
Cos ϕ_1 (70 km)	0.83	0.75	0.60	0.38	0.22	0.17	0.16	0.16
K	98.2	98.2	98.2	98.2	98.2	98.2	98.2	98.2
P_r	14.6	14.6	14.6	14.6	14.6	14.6	14.6	14.6
$-D_p$	−34.0	−37.0	−40.0	−46.0	−52.0	−57.0	−60.0	−62.0
A_t	−5.0	−3.0	−1.9	−0.7	−0.2	−0.1	0	0
$-L_t$	0	0	0	−1.0	−4.0	−12.0	−19.0	−22.0
$-I$	−32.7	−32.0	−30.3	−27.0	−23.4	−21.5	−21.0	−21.0
C	0.2	0.3	0.4	0.6	1.7	5.7	6.2	7.2
$-L_r$	6	6.0	6.0	6.0	6.0	4.5	2.6	2.0
A_r	−5.0	−3.0	−1.9	−0.7	−0.2	−0.1	0	0
H	0	0	0	0	0	0	0	0
E_{s1}	42.3	44.1	45.1	44	40.7	33.3	21.6	17.0

TABLE 2. *Skywave field intensity calculation—Thule to Goose Path*

$E_{s1}=K+P_r-D_p+A_t-L_t-I+C_1-L_r+A_r$

Condition: Cos $\chi \approx 0.2$, ionosphere low (70 km) with ionization attenuation typical of nighttime conditions.

Distance nautical miles	70	100	200	400	700	1,000	1,250	1,400
ψ_1(70 km)	47°	37°	20°	9°	3.2°	0.3°	−0.5°	−2°
Cos ϕ_1(70 km)	0.75	0.60	0.38	0.22	0.17	0.16	0.16	0.15
K	98.2	98.2	98.2	98.2	98.2	98.2	98.2	98.2
P_r	14.6	14.6	14.6	14.6	14.6	14.6	14.6	14.6
$-D_p$	−37.0	−40.0	−46.0	−52.0	−57.0	−60.0	−62.0	−63.0
A_t	−3.0	−1.9	−0.7	−0.2	−0.1	0	0	0
$-L_t$	0	0	−1.0	−4.0	−12.0	−19.0	−20.0	[a] −20.0
$-I$	−11.8	−9.8	−7.6	−5.0	−3.8	−3.6	−3.6	−3.6
C_1	0.3	0.4	0.6	1.7	5.7	6.2	7.2	8.6
$-L_r$	6.0	6.0	6.0	6.0	4.5	2.6	2.0	[b] −12.0
A_r	−3.0	−1.9	−0.7	−0.2	−0.1	0	0	0
E_{s1}	64.3	65.6	63.4	59.1	51.0	39.0	36.4	22.8

[a] This low value for L_t is due to ψ being negative and the close proximity of the sea in the launching path.
[b] The change in L_r is due to the crossing of the sea coast at 1,250. Thus at 1,400 the receiving antenna launching area was over arctic soil.

TABLE 3. *Skywave field intensity calculation—Thule to Goose Path*

$E_{s2}=K+P_r-D_p+A_t-L_t-I_1+C_2-I_2-G-L_r+A_r$

Condition Cos $\chi \approx 0.2$, ionosphere low (70 km) with ionization attenuation typical of nighttime conditions.

Distance nautical miles	200 [a]	400 [a]	700 [a]	1,000	1,250	1,400
ψ_2(70 km)	36°	20°	11°	6°	4°	3°
Cos ϕ_2(70 km)	0.6	0.37	0.24	0.18	0.17	0.17
K	98.2	98.2	98.2	98.2	98.2	98.2
P_r	14.6	14.6	14.6	14.6	14.6	14.6
$-D_p$	−48.0	−53.0	−57.0	−60.0	−62.0	−63.0
A_t	−1.9	−0.7	−0.2	−0.1	−0.1	0
$-L_t$	0	−1.0	−4.0	−7.0	−10.0	−12.0
$-I_1$	−10.0	−7.5	−5.0	−4.0	−4.0	−4.0
C_2	0.2	0.4	1.5	2.0	3.5	5.0
$-I_2$	−10.0	−7.5	−5.0	−4.0	−4.0	−6.0
$-G$	0	0	0	0	[b] −3.0	[b] −3.0
$-L_r$	6.0	6.0	6.0	5.5	5.0	[c] −2.5
A_r	−1.9	−0.7	−0.2	−0.1	−0.1	0
E_{s2}	47.2	48.8	48.9	45.1	38.1	27.3

[a] Whenever the sky ray distance (D_s) is appreciably different from the surface distance (D), see figure 16 to calculate D_s and use this value to calculate D_p.
[b] At these distances the ground reflection point was on Baffin Island.
[c] At this point the receiving antenna was over land instead of sea as it had been up to 1,250 nautical miles.

and we chose a typical daytime ionospheric height of 70 km with ionospheric reflection values for normal nighttime conditions. Conditions for flight 2 were considered to be normal for summer daytime conditions in northern latitudes. Since cos $\chi \approx 0.2$ at these latitudes during the summer, we used ionospheric reflection values close to what we have indicated on figure 22 as typical values for winter noon conditions.

The authors acknowledge the assistance received from K. A. Norton, J. R. Wait, and J. S. Belrose during many helpful discussions of these problems as well as for the permission to use material prepared by them. In addition, they are indebted to: Col. W. E. Geyser of AACS for his interest and assistance in making this work possible, and the personnel under his command including O. E. Cummings who made the aircraft measurements of flight 2; R. M. Coon and E. F. Florman for helpful discussions and assistance in obtaining the data; Winifred Mau for her assistance during the preparation of this manuscript; J. C. Harman for preparation of the illustrations; and to W. W. Brown for permission to use some of his vlf observations. The cooperation of J. E. Gillis, A. B. Nicolaou, and Hans Roethlisberger of the Snow, Ice, and Permafrost Research Establishment, and Comdr. C. Santae, USN, is also appreciated.

4. References

[1] J. R. Wait, Transmission curves for ground-wave propagation at low radio frequencies, Radio Physics Lab., Defence Research Telecommunications Establishment, Ottawa, Canada, rept. No. R–1, April, 1953.

[2] G. Millington, Ground-wave propagation over an inhomogeneous smooth earth, Proc. Inst. Elec. Engrs. London **96 Pt. III,** 53 (1949).

[3] J. R. Wait, Mixed-path ground-wave propagation, I and II, J. Research NBS **57,** 1 (1956) RP 2687; and **59,** 19 (1957) RP 2770.

[4] J. R. Wait, The amplitude and phase of the low-frequency ground wave near a coastline, J. Research NBS **58,** 237 (1957) RP2756.

[5] J. R. Wait and A. Murphy, The influence of a ridge on the low-frequency ground wave, J. Research NBS **58,** 1 (1957) RP2727.

[6] F. E. Terman, Radio Engineers Handbook, sec. 11 (McGraw-Hill Book Co., New York, N.Y., 1943).

[7] J. R. Wait and W. A. Pope, Input resistance of L.F. unipole aerials, Wireless Eng. **32,** 131 (1955).

[8] K. A. Norton, The calculation of ground-wave field intensities over a finitely conducting spherical earth, Proc. I.R.E. **29,** 623 (1941).

[9] N. E. Dorsey, Properties of ordinary water substance (Reinhold Publishing Corp., New York, N.Y., 1940).

[10] J. R. Wait and A. Conda, Trans. I.R.E., **AP-6** (October 1958).

[11] K. A. Norton and A. C. Ombert, The maximum range of a radar set, Proc. I.R.E. **35,** 4 (1947).

[12] K. A. Norton, (Informal communication, July 1958).

[13] R. N. Bracewell, K. G. Budden, J. A. Ratcliffe, T. W. Straker, and K. Weekes, The ionospheric propagation of low- and very-low-frequency radio waves over distances less than 1000 km, Proc. Inst. Elec. Engrs. London **98 Pt. III,** 221 (1951).

[14] A. H. Waynick, The present state of knowledge concerning the lower ionosphere, Proc. I.R.E. **45,** 741 (1957).

[15] J. E. Bickel, J. L. Heritage, and S. Weisbrod, An experimental measurement of vlf field strength as a function of distance, using an aircraft, U.S. Navy Electronics Lab., San Diego, Calif., rept. No. 767, Jan. 28, 1957.

[16] J. S. Belrose, Some investigations of the lowest ionosphere (Doctoral Thesis, University of Cambridge, 1958).

[17] J. E. Bickel, A method for obtaining lf oblique-incidence reflection coefficients and its application to 135.6 kc/s data in the Alaskan area, J. Geophys. Research **62,** 373 (1957).

[18] Rune Lindquist, Preliminary results of an investigation of propagation conditions at very low frequencies, Union Radio Scientifique Internationale, Boulder, XII General Assembly Document, 1957/No. 104/Commission III.

[19] Weekes and Stuart, The ionospheric propagation of radio waves with frequencies near 100 kc/s over distances up to 1000 km, Proc. Inst. Elec. Engrs. London **99 Pt. IV,** 38 (1952).

[20] S. A. Bowhill, The fading of radio waves of frequencies between 16 and 2400 kc/s, J. Atmospheric and Terrest. Phys. **8,** 129 (1956).

[21] J. R. Wait and L. B. Perry, Calculations of ionospheric reflection coefficients at very-low radio frequencies, J. Geophys. Research **62,** 43 (1957).

[22] J. R. Wait, The mode theory of vlf ionospheric propagation for finite ground conductivity, Proc. I.R.E. **45,** 760 (1957).

[23] J. R. Pierce, Sky-wave field intensity I. Low- and very-low radio frequencies, Cruft Lab. tech. rept. 158 (Harvard U., Cambridge, Mass., Sept. 1, 1952).

[24] A. D. Watt, R. M. Coon, E. L. Maxwell, and R. W. Plush (informal communication, 1957).

[25] B. Landmark, The fading of radio waves reflected from the E layer, J. Atmospheric and Terrest. Phys. **10,** 288 (1957).

Boulder, Colo. (Paper 63D1–12).

Selected Abstracts of Publications by the National Bureau of Standards Staff

Bureau Publications

Propagation of very-low-frequency pulses to great distances, J. R. Wait, *J. Research NBS* **61,** *187–203 (Sept. 1958) RP2898.*

A theoretical study is presented for the propagation of electromagnetic pulses at very low frequencies to large distances. The space between the earth and the ionosphere is represented as a wave guide with sharply bounded and concentric spherical boundaries. The concept of phase and group velocity and its application to the present problem is discussed in some detail. The influence of the propagation medium on the shape of the envelope of a quasi-monochromatic pulse is also considered. Using an alternative approach, the response of an impulsive source is also calculated and is shown to be a damped oscillatory function of time with a quasi-half-period varying in a predictable manner with distance of travel in agreement with the observations of Norinder and Hepburn.

Techniques for accurate measurement of antenna gains, H. V. Cottony, *NBS Circ. 598, 10 pp. (Dec. 1958).*

Comparison of published results of experimental antenna measurements, particularly gains, reveals apparent discrepancies of the order of one or more decibels. Experimental work at the National Bureau of Standards on sealed model antennas for long-range VHF communication via ionospheric scatter revealed some sources of difficulties and led to the adoption of special precautions resulting in significantly more consistent, and, it is believed, more accurate results. The procedures are based on the comparison method but include, in addition to the standards antenna, the use of a third antenna here designated as the reference antenna. To obtain more accurate measurements it was found essential to correct for the standing wave pattern in the field set up, presumably, by reflections from the irregularities in the terrain. Special features of instrumentation, including methods for minimizing and measuring matching losses, are described. The accuracy of the techniques has been verified by measuring the gain of an antenna, the value of which could be accurately calculated.

Transmission and reflection of electromagnetic waves in the presence of stratified media, J. R. Wait, *J. Research NBS* **61,** *205–232 (Sept. 1958) RP2899.*

A general analysis is presented for the electromagnetic response of a plane stratified medium consisting of any number of parallel homogeneous layers. The solution is first developed for plane-wave incidence and then generalized to both cylindrical and spherical-wave incidence. Numerical results for interesting special cases are presented and discussed. The application of the results to surface-wave propagation over a stratified ground is considered in some detail.

Non-Bureau Publications

Antennas and waveguides, and annotated bibliography (169 refs), U.S.A. National Committee Report (URSI Subcommission 6.3), H. V. Cottony, R. S. Elliott, E. C. Jordan, V. H. Rumsey, K. M. Siegel, J. R. Wait, and O. C. Woodyard, *IRE Trans.* **AP–7,** *87–98 (Jan. 1959).*

Recent developments in some of the more active areas of antenna and diffraction theory are summarized. Although concerned primarily with U.S.A. activity, some reference is made to the literature of other countries where closely allied work is in progress. Fields included in this survey are broadband antennas, wide-angle microwave optics, antennas for ionospheric scatter propagation, traveling-wave antennas, slot radiators, diffraction, and scatter theory.

Calculations of transverse current loss in buried wire ground systems, J. R. Wait, *Appl. Sci. Res. Sect. B* **7,** *81–86 (1958).*

A formula is developed for the impedance between a wire grid buried in a homogeneous ground and an overhead conducting plane. The result can be employed to calculate the power lost due to transverse currents in buried wire ground systems for antennae.

Characteristic size of airglow cells, F. E. Roach, E. Tandberg-Hanssen, and L. R. Megill, *J. Atmospheric and Terrest. Phys.* **13,** *113–121 (Dec. 1958).*

Evidence is presented for the existence of discrete 5577 airglow cells in the upper atmosphere. Independent analyses are based on a) the comparison of diurnal intensity variations in different parts of the sky and b) the measurement of intensity gradients on airglow isophote maps. The numerical estimation of cell size is complicated by the fact that a single cell is, in general, larger than the region observable at a given location. The typical airglow cell for the 12 nights of the study has a diameter of approximately 2500 km.

Compact microwave refractometer for use in small aircraft, M. C. Thompson and M. J. Vetter, *Rev. Sci. Instr.* **29,** *1093–1096 (Dec. 1958).*

A microwave refractometer has been developed for use in a Type L–20 "Beaver" single-engine light aircraft. A two-channel recording system is used providing simultaneous profile and turbulence records with rms noise levels of approximately 0.05 and 0.01 N unit, respectively. Power requirement is 25 amperes at 28 volts dc, and installation requires one-half hour for two men. No structural modification of the aircraft is required. The results of flight tests are presented.

Day-to-day coordination of IGY observations, A. H. Shapley, *Proc. IRE* **47,** *323–327 (Feb. 1959).*

The IGY World Days and Communications program has four facets: (1) an IGY Calendar of selected days for experiments which cannot be carried on continuously; (2) a scheme for specifying periods when phenomena of unusual interest are anticipated a few hours in advance; (3) the prompt and wide distribution of summary observations of outstanding solar and geophysical events; and (4) the use of the world-wide scientific communications network for processing information of temporal interest. This is a cooperative effort including all the countries participating in the IGY. The paper describes the plan and its execution, including the arrangements for rapid communications.

Excitation of surface waves on plane and curved impedance boundaries, J. R. Wait, *Congrès International sur La Propagation des Ondes Radio-Electriques, Liège, Oct. 6–11, Résumé des communications, 26 1958).*

The excitation of surface waves on conducting, stratified, dielectric clad and corrugated surfaces has been studied in some detail. An expression for the field of an electric dipole located over a flat surface with a specified surface impedance Z was derived from the formal integral solution by a modified saddle-point method. Using the value of Z appropriate for a homogeneous conducting ground the general expressions reduce to those given by Norton. In this case, the phase of Z lies between 0 and 45°. When the phase exceeds 45° such as it may for a stratified ground, the radiated wave of the dipole becomes partially trapped to the interface. This effect is most pronounced for an inductive surface where the phase of Z is 90° in which case the energy of the wave is confined to within a small distance from the surface. Such

inductive surfaces are a metallic plane with a thin dielectric film or a corrugated surface. This unifying treatment provides a link between the surface waves of Zenneck, Sommerfeld, Norton and Goubau and indicates that the magnitude and phase angle of Z controls to what extent these waves may exist for a dipole excitation. The analysis has been extended to curved surfaces possessing a reactive surface impedance. The model chosen is a magnetic line or slot source on the surface of circular cylinders with an inductive or capacitive boundary condition. Spherical, elliptical and spheroidal surfaces have also been considered. The radiation pattern of such systems has been calculated using three methods, whose range of validity overlap.

Gains of finite-size corner-reflector antennas, H. V. Cottony and A. C. Wilson, *IRE Trans.* **AP–6,** *366–369* (*Oct. 1958*).

An experimental corner reflector was erected at the Table Mesa antenna range near Boulder. The aperture angle of this antenna was made adjustable to any value between 20 and 180 degrees. The widths and lengths of the reflecting surfaces were each adjustable from 0.4 to 5.0 wavelengths. Measurements of gain were made for numerous combinations of lengths and widths of reflecting surfaces. These measurements were made with a half-wave dipole in the first, second and third maximum positions. The aperture angle was adjusted to maximize the gain. The principal results are presented in the form of contours of constant gain plotted for a range of widths and lengths of reflecting surfaces from 0.4 to 5.0 wavelengths. These graphs should be useful to a designer of corner-reflector antennas.

Induction by an oscillating magnetic dipole over a two-layer ground, J. R. Wait, *Appl. Sci. Res. Sect. B* **7,** *73–80* (*1958*).

An expression for the mutual electromagnetic coupling between two small loops over a two-layer ground is derived. The result is expressed in a form which is suitable for calculation by a digital computer. When the heights or separation of the loops are large compared to the skin depth in the ground, simple asymptotic formulas for the fields can be developed. The results are employed to obtain a convenient formula for the self-impedance of a loop over a two-layer ground.

Interpretation of night-time low-frequency ionograms, J. M. Watts, *J. Geophys. Research* **63,** *717–726* (*Dec. 1958*).

Two types of retardation of the extraordinary ray on passing through a region whose critical frequency is below the gyrofrequency are discussed. On a sweep-frequency basis, one is observed near the critical frequency of the ordinary ray and the other appears near the gryo-frequency. Observations of either require near transparency of the lower E-region during the night, but information as to the total number and distribution of electrons below the F-layer can be deduced from them.

Long distance pulse-propagation experiment on 20.1 megacycles, R. Silberstein, *J. Geophys. Research* **63,** *445–466* (*Sept. 1958*).

A pulse-propagation experiment was performed between Sterling, Virginia, and Maui, Territory of Hawaii, with the object of studying the mechanism determining the classical MUF at long distances, and also of obtaining a general idea of the mode structure. Simultaneous oblique-incidence records of the transmitter pulses were made at Boulder, Colorado which lies along the path. Some back-scatter records were obtained at the transmitter site; and vertical-incidence records were made at the midpoint of the Sterling to Boulder portion of the path. Results showed greatly differing mode structures between one day and the next, indicating that the long path is very sensitive to ionospheric conditions, and also that M and N reflection and layer tilts play an important part. Experimental evidence was strong that sporadic-E ionization enables the propagation of a wave over the path where F2 ionization at one end of the path is not sufficient for reflection.

Low and medium frequency radio propagation (abstract), K. A. Norton, *Congres International sur la Propagation des Ondes Radio-Electriques, Liège, Oct. 6–11, Résumé des communications, 20* (*1958*).

During recent years, extensive progress has been made in understanding the nature of ionospheric propagation in the frequency range from 30 to 1,000 kc. In particular, Wait and Conda have developed suitable methods for determining the effect of the curvature of the earth on the illumination of the ionosphere by antennas radiating at angles near and below grazing incidence. Belrose has developed empirical formulas for ionospheric absorption in the range 70 to 250 kc. Watt, Maxwell, and Whelan have analyzed data which makes possible the extension of these formulas to still lower frequencies, and Bean and Thayer have developed a radio standard tropospheric atmosphere which made possible good estimates of the bending of the radio waves in propagation to and from the ionosphere. Making use of these results, together with earlier studies made by the author of the absorption of radio waves in the band 540 to 1,600 kc and some new results on focusing at the curved surface of a rough ionosphere, predictions are made of the transmission loss expected in ionospheric propagation by m reflections at the ionosphere. The predictions appear to be in general agreement with the available experimental data, although the paper clearly points up the desirability of more definitive studies of the influence of absorption and polarization. The methods in that paper indicate that the transmission loss increases only very slowly as the angle of departure approaches and goes well below grazing incidence with the earth. This provides an explanation for recent experimental data obtained with a pulse system as reported by Doherty which indicates very large intensities for one-hop sky waves on 100 kc both day and night at ranges up to 1,800 miles.

Measurement of ground conductivity at VLF, J. R. Wait and A. M. Conda, *IRE Trans.* **AP–6,** *273–277* (*July 1958*).

The applicability of the four electrode methods of measuring ground conductivity at very low radio frequencies is discussed. The general theory is extended to include anisotopy in the substrata. In view of the spurious coupling between the current and potential line in conventional configurations, an alternative array is proposed which is arranged so that the inductive coupling is zero. A number of charts are computed which facilitate the interpretation of the measured or apparent conductivity in terms of a two-layer earth.

Movements of airglow cells, F. E. Roach, E. Tandberg-Hanssen, and L. R. Megill, *J. Atmospheric and Terrest. Phys.* **13,** *122–130* (*Dec. 1958*).

The possibility of translatory and rotatory motions of airglow cells is examined. Evidence is presented for translatory motions in the 100 km region with speeds of the order of 100 m/sec. Rotatory motions are indicated corresponding to a mean period of 5 hr. A broad similarity seems to exist between airglow cell motions in the 100 km region and vortex cells near sea level. The problem of the mechanism responsible for the excitation of the oxygen atom to the S state is discussed with particular reference to (a) photochemical reactions and (b) environmental effects.

Night airglow, Franklin E. Roach, *Proc. IRE* **47,** *267–271* (*Feb. 1959*).

A phenomenological description of the night airglow is presented, reviewing the historical background, and what is known about height, temporal and spatial variations in intensity, and movements. The very important relationship to aurora and evidence for latitude-seasonal effects are examined.

Pattern of an antenna on a curved lossy surface, J. R. Wait and A. M. Conda, *IRE Trans.* **AP–6,** *348–359 (Oct. 1958).*

Extensive numerical results are presented for the radiation fields of electric and magnetic type antennas mounted on smooth curved surfaces of finite conductivity. The model chosen is a circular cylinder whose surface impedance is specified. A residue series representation is employed for the portion of space deep in the shadow while a geometrical-optical representation is used in the "lit" region. In the penumbra, the fields are expressed in terns of the "Fock functions." The results are also applicable to other smoothly varying curved surfaces such as spheres, parabolic cylinders, and paraboloids. As an application, the E-plane patterns are computed for a small loop antenna on a spherical earth for both sea and land illustrating the so-called cut-back effect.

Patterns of a slot-array antenna on a finite and imperfect ground plane, J. R. Wait and A. M. Conda, *L'Onde Electrique (Proc. International Congress Ultra High Frequency Circuits and Antennas, Paris, Oct. 21–26, 1957)* **38,** *21 (1958).*

The radiation from a slot aperture on a perfectly conducting half-plane is discussed in some detail. The extension to end-fire arrays of slots is also treated. Numerical computations are carried out for radiation patterns. It is shown that in general, the main beam is tilted up and away from the edge of the half-plane. The situation where the half-plane is lying on the surface of a homogeneous flat earth is also considered. The results have application in the design of flush mounted antennas of the end-fire type where, because of practical limitations, the ground plane is of finite extent and may be located on the surface of a "lossy" plane earth.

Propagation of a ground wave pulse around a finitely conducting spherical earth from a damped sinusoidal source current, J. R. Johler and L. C. Walters, *IRE Trans.* **AP–7,** *1–10 (Jan. 1959).*

The form of the transient electromagnetic ground wave which has been propagated over a finitely conducting spherical earth from a source current dipole is calculated by a direct quadrature evaluation of the Fourier integral.

Radio propagation research at the National Bureau of Standards, F. W. Brown, *Am. J. Phys.* **26,** *628–634 (Dec. 1958).*

The Central Radio Propagation Laboratory is the central agency of the U.S. Government for obtaining, analyzing, and disseminating information on the propagation of radio waves at all frequencies along the surface of the earth, in the atmosphere, and in the space. In order to carry out its responsibility, this Laboratory conducts research on the nature of radio waves, the nature of the media through which radio waves are transmitted, the interaction of the waves with the media, and the nature of radio noise and interference effects. This article describes some of that research and summarizes the historical background leading to the present framework for carrying out the research. To the best of the author's knowledge, no completely comprehensive general summary of the Laboratory's history, objectives, and research programs has previously been published.

Reduction of adjacent-channel interference components from frequency-shift-keyed carriers, A. D. Watt, V. J. Zurick, and R. M. Coon, *IRE Trans.* **CS–6,** *39–37 (Dec. 1958).*

The abrupt changes in frequency associated with binary modulation generate sideband components which frequently interfere with services in adjacent channels. Reductions in interfering bandwidths of 10 to 1 or greater over unfiltered keying at sideband levels 80 db or more below the unmodulated carrier are possible by the use of relatively simple filters in the keying circuit of frequency-shift-keyed transmitters. The frequency spectra resulting from filtered keying waves of FSK transmitters are derived, and the results presented in graphic form. The sizeable reduction of the interfering sideband components as predicted by the analysis is found to agree very closely with experimental measurements. Characteristics of desirable keying circuit filters to obtain these reductions are specified.

Riometer—a device for the continuous measurement of ionospheric absorption, C. G. Little and H. Leinbach, *Proc. IRE* **47,** *315–320 (Feb. 1959).*

A sensitive, self-balancing, noise-measuring equipment, known as the riometer, is described. This instrument has been designed for routine measurement of ionospheric absorption during the IGY, using the cosmic noise method. Application of this technique in the auroral zone has resulted in quantitative measurements of ionospheric absorption, even during polar blackouts. The riometer has the advantages over a simple total-power cosmic noise receiving system of (1) linear response to changes of input noise power, (2) high accuracy in the presence of narrow-band RF interference, and (3) good long-term stability.

Solar cycle influence on the lower ionosphere and on VHF forward scatter, C. Ellyett and H. Leighton, *Proc. IRE* **46,** *1711–1716 (Oct. 1958).*

A brief survey of published literature indicates that the electron concentration in the D region of the ionosphere changes with the solar cycle. Such changes should also be apparent in VHF forward scatter signal intensities. Analyses have therefore been made of seven years' results, at 49.8 mc from the Cedar Rapids, Iowa, to Sterling, Va., forward monthly median received signal intensity, for the noon and afternoon period only, follows the same trend as the mean monthly sunspot number. The effect is still more pronounced when comparison is made with magnetic disturbance indexes. However, when the comparison with magnetic indexes is made using a direct comparison in each 3-hour period, and eliminating long-term trends, no observable increase of the signal intensity is found as the magnetic conditions become disturbed. This is in sharp contrast to the behavior of such circuits in auroral regions. Further analysis shows that a weak increase in signal intensity with a rising magnetic index does occur if the analysis is confined to the midday period. It is concluded, therefore, that any magnetic influence is smaller on middle than on high latitude scatter circuits, but an effect can be clearly discerned by studying the magnitude of the received signal through a solar cycle. These results show that VHF forward scatter signals are not caused by reflection from meteor trails alone.

Transmission loss curves for propagation at very low radio frequencies, J. R. Wait, *IRE Trans.* **CS–6,** *58–62 (Dec. 1958).*

Curves of the transmission loss are presented for the propagation to great distances at frequencies in the range 10 to 20 kc. The theoretical model of the ionosphere assumed is a sharply bounded homogeneous ionized medium. The working formula for the field is a sum of waveguide-type modes. The calculated results compare favorably with experimental data at 16.6 kc over the Pacific Ocean.

Very-low-frequency emissions generated in the earth's exosphere, Roger M. Gallet, *Proc. IRE* **47,** *211–231 (Feb. 1959).*

Naturally occurring, very low-frequency signals not associated with lightning discharges, and strongly correlated with solar activity, have been recognized nearly as long as the atmospheric whistlers which have their genesis in lightning discharges. Whereas whistlers have been satisfactorily explained, until recently these other phenomena have not. From the examination of a large quantity of high resolution spectrograms, it has been deduced that a major fraction, if not all, of these other "noises" are excited in the exosphere by streams and bunches of high-speed ionized particles precipitating into the ionized atmosphere in the presence of the earth's magnetic field. The electromagnetic waves excited are then propagated in the manner of whistlers.

List of Other Publications by the National Bureau of Standards Staff [1]

Atmospheric bending of radio waves (abstract), B. R. Bean, *Congres International sur la Propagation des Ondes Radio-Electriques, Liège, Oct. 6–11, Résumé des communications, 3 (1958).*

Communications via meteor bursts, G. F. Montgomery, *Radio Electronics,* **29,** *88–90 (July 1958).*

Comparison of phase difference and Doppler shift measurements for studying ionospheric fine structure using earth satellites (letter), M. C. Thompson and D. M. Waters, *Proc. IRE* **46,** *1960 (Dec. 1958).*

First meeting on radio climatology (letter), B. R. Bean, *Proc. IRE* **46,** *1425–1426 (July 1958).*

Mean deviation of the Poisson distribution, E. L. Crow, *Biometrika* **45,** *556–559 (Dec. 1958).*

Observation of vertical-incidence scatter from the ionosphere at 41 Mc/sec., K. L. Bowles, *Phys. Review Letters* **1,** *454–455 (Dec. 1, 1958).*

Preface to proceedings of the third symposium on cosmical gas dynamics, R. N. Thomas and J. M. Burgers, *Held at the Smithsonian Astrophysical Observatory, Cambridge, Mass. June 24–29, 1957, Review of Modern Physics* **30,** *908–911 (July 1958).*

Properties of lightning impulses which produce whistlers (letter), R. A. Helliwell, A. G. Jean, and W. L. Taylor, *Proc. IRE* **46,** *1760–1762 (Oct. 1958).*

Property of additively closed families of distributions, E. L. Crow, *Ann. Math. Stat.* **29,** *892–897 (Sept. 1958).*

Proposed recommendation on ionospheric vertical soundings after the IGY, A. H. Shapley (*Memorandum No. 18*) *URSI (International Scientific Radio Union), Inform. Bull. No. 111, 42 (1958).*

Recent theoretical studies of the propagation of long waves (abstract) J. R. Wait, *Congrès International sur la Propagation des Ondes Radio-Electriques, Liège, Oct. 6–11, Résumé des communications, 27 (1958).*

Study of earth currents near a VLF monopole antenna with a radial wire ground system (letter), J. R. Wait, *Proc. IRE* **46,** *1539–41 (Aug. 1958).*

Transmission error function for meteorburst communication (letter), G. Franklin Montgomery, *Proc. IRE* **46,** *1423–1424 (July 1958).*

Tropospheric effects on 6-Mc pulses (letter), R. Silberstein, *Proc. IRE* **46,** *1968 (Dec. 1958).*

World-wide soundings committee, A. H. Shapley, *Memorandum 17–Future work, URSI Inform. Bull. No. 110, 55 (1958).*

---

[1] After this initial issue an attempt will be made to include everything published during the previous two month period.

U. S. GOVERNMENT PRINTING OFFICE: 1959

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September-October 1959

Stratification in the Lower Ionosphere[1]

C. Ellyett[2] and J. M. Watts

(April 16, 1959)

A survey of the evidence for stratification in the ionosphere below 100 km is given, covering radio and optical observations, and rocket measurements. The conclusion is reached that one stratum at about 85 km is observed consistently, and that other fine structure exists but has no long-time constancy of height or pattern. There is no series of preferred heights below 100 km. The authors consider explanations which may account for the observations, and advocate the testing of radio methods of exploration in conjunction with rocket measurements in order to develop the most practicable means of obtaining accurate electron density versus height profiles on a synoptic basis.

1. Introduction

Gradually, over the past two decades, experimental results have either indicated directly, or have required the postulation of the existence of one or more discrete horizontal layers of ionization below the *E*-region of the ionosphere. Some of the most direct experimental evidence for such stratification has been obtained from vertical pulsed sounding of the ionosphere by Dieminger (1952), Gardner and Pawsey (1953), Gregory (1956), and others. However, many other unrelated types of experiment in the literature tend to support the concept of stratification, and, because of the diversity of experiments and the types of phenomena observed, the term "stratification" has been broadened in this paper to include, in addition to layers of ionization, any phenomenon tending to localize or show maxima at one or more heights, provided the effect is not purely transient in character.

The purpose of this paper is threefold: First, to search the literature for all references to discrete heights of phenomena occurring in the *E*-region and below; secondly, to present and compare models based on two different premises; and, finally, to discuss the avenues most likely to lead to further progress in exploring this region of the atmosphere between 50 and 110 km.

The references to discrete heights have been surveyed and grouped under the headings of radio, rocket, meteor, and optical phenomena, and are followed by a briefer study of the heights of occurrence of some of the relevant photoemissive processes. The existence of winds and turbulent air conditions in the lower ionosphere is well established and heights of these phenomena have in some cases been deduced. Sufficient material has been included to show the nature of the phenomena, but a comprehensive discussion of turbulence has not been attempted.

2. Radio Methods

2.1. Conventional Soundings

At first, workers were busy discovering the important characteristics of the *E*- and *F*-regions above 100 km, and paid scant attention to less prominent details. Appleton (1930), using a frequency change method of investigating the standing wave pattern of sky and ground waves, noticed "subsidiary fringes" which could be interpreted as the effect of extra layers in the *E*-region. Ratcliffe and White (1933) showed that some dual echoes which occurred in the *E*-region but never in the *F*-region were not magnetoionic components and were presumably from extra stratification. Both of these cases were probably from heights just above 100 km, and would now be called sporadic *E*, but Appleton and Piddington (1938) produced a histogram of echo heights showing a broad peak at 110 km, but significant occurrences of heights down to 80 km. In their opinion all echoes below the *E*-layer were transient in character and never of high strength. Appleton, Naismith, and Ingram (1939) discussed the application and limitations of the critical frequency method of determining maximum electron densities of the layers, and in so doing, mentioned the effects of thin layers, scattering centers imbedded in layers, thin sheets formed by coagulation of clouds of electrons, 90- to 100-km echoes (called trailing echoes), and *F*-region scatter. They also showed that sporadic-*E* reflections are from a preferred height. Ellyett (1947) collected early *D*-region information, and presented it together with some Pacific Island data from Pitcairn and Raoul Islands, showing that *D*-echoes were observed only during the day. Whale (1951), in summarizing the qualitative interpretation of regular ionograms, discussed subsidiary layers and evidence of irregularities in the *E*-region.

[1] This work was carried out in part under Contract AF 64(500)-5 from the U.S. Air Force Cambridge Research Center.

[2] University of Canterbury, Christchurch, New Zealand.

2.2. Experiments With Pulses at LF and MF

Pulse transmissions at low and medium frequencies have produced a large quantity of data concerning the lower ionosphere. Helliwell (1949) and Helliwell, Mallinkrodt, and Kruse (1951) observed at 100 and 325 kc, reporting heights from 86 to 106 km which were constant over periods of several hours, but which varied widely from day to day. At night "multiplicity of strata between 90 and 130 km" was observed. The separation of the echoes did not increase with order, which they interpreted, in the manner of Ratcliffe and White (1933), to mean stratification of the medium instead of echoes magnetoionically split. Helliwell (1952) also has shown that when histograms of observed reflection heights are plotted, certain "preferred heights" are evident. During August 1951 the values obtained were 100, 106, and 112 km.

Further information on observed heights of reflection may be found in Lindquist (1953) and Watts and Brown (1954). Lindquist reported vertical incidence results at 150 kc in the form of monthly median values of reflection heights. They were generally 85 km by day and 95 km by night. Watts and Brown produced low frequency ionograms with sweep frequency equipment, and reported heights in the forms of histograms and median values for 4 winter months. Day levels were from 70 to over 100 km, and at night the lowest level was 90 km. Several discrete levels were observed day and night, but in addition there was some evidence of turbulence in the lower reflections.

Kilpatrick (1957) has correlated low frequency ionograms with standard high frequency ionograms made near the same location. The results clearly indicate that some of the low frequency reflections have little or no relation to the normal daytime *E*-layer. In order to investigate the reliability of navigational systems—especially loran—Naismith and Bramley (1951) conducted a series of experiments using pulses from regular loran stations, modified broadcast transmitters on 668 and 877 kc, and regular vertical incidence sounders. Equivalent heights of reflection were computed for different ranges and days. They found a marked failure of Martyn's theorem concerning reflection delays at vertical and oblique incidence, indicating that partial reflections were important at oblique incidence. The lowest heights derived were 50 km at long range (1,208 km) to 85 km at 170 km. Two peaks, one at 75 km and another at 95 km were commonly present at intermediate ranges. At a range of 1,208 km the only prominent peak in the histogram was at 70 km.

In another oblique incidence experiment at 300 kc (Watts 1952 a), derived reflection heights were 70 km by day and 80 km by night. It is significant that the oblique incidence results gave lower heights than had been observed at vertical incidence at any frequency. Also when 2- and 3-hop modes were used to derive reflection heights, greater values always resulted, showing that the angle of incidence controlled the effective height of reflection.

Polarization tests of returned echoes have been made in order to resolve the question of the meaning of the apparent stratification in the region below 100 km. Benner, Grace, and Kelso (1950) used a crossed-loop polarimeter receiver to determine whether the split echoes were due to magnetoionic splitting, layer stratification, or partial reflection. They found the measured polarization precluded magnetoionic splitting, but it did not fit their theoretical limiting polarization. Watts (1952 b) used a circularly polarized transmitting antenna, with periodic switching of the sense of rotation. Most occasions showed that split echoes had appreciably different polarizations, as evidenced by different amplitudes when the sense was reversed, but in only one case were the split echoes of completely opposite sense.

In the low level region, disturbances by magnetic storms have the effect of making the region more transparent at night to the low and medium frequencies. Also, the fading is more rapid (Watt and Brown, 1951, and Lindquist, 1953). Lindquist also noted that SID's reduced the amplitude of the 150-kc reflections, but did not change their virtual height. This probably means that fadeouts are caused by an increase of ionization well below the reflection heights being observed.

Fejer (1955) used a pulse technique of novel type in an experiment involving interaction of two pulses in the ionosphere. By adjusting the timing of the two trains of pulses, the height at which they coincided could be altered. The method enables a determination of both electron density and collision frequency versus height to be made. He derived a profile of electron density showing ledges near 72 and 78 km, presumably in the daytime. The method seems potentially powerful but, like most interaction techniques, requires dealing with very small effects.

2.3. Experiments Involving VLF Using Either CW or Sferics

An approach to lower ionospheric understanding by means of long wave propagation has interested some other workers. It was hoped that the experiments would prove that a simple model would suffice to predict propagation characteristics.

Weekes (1950) at 16 kc found that the reflecting layer changed its characteristics between 300 and 500 km distance from the transmitter. At less than 500 km the received polarization was circular and the reflection height was either 67 or 75 km, depending on alternative phase changes assumed upon reflection. At 500 km or greater the polarization was linear, but the height was not determined. Bain and Bracewell (1952) proposed that two different reflective heights, having different diurnal characteristics, would explain the VLF phenomena. Bracewell (1952) postulated a new theory of layer formation to explain the downward extension of ionization observed during an SID, and also that there must be a separate layer below the *E*-layer. Bain (1953) went further, requiring two layers below the *E*, and re-

quiring that the SID effect was a downward extension of the top layer without change in the nature of its ionizing reaction. Hopkins and Reynolds (1954) still later considered the problem, and decided that difficulties of interpretation arise when more than one wave of appreciable magnitude is received and that their results gave only a broad description of oblique LF and VLF propagation.

Rivault (1950) observed that the average height of reflection of sferics in October, 1949, from 1700 to 2100 hr was 75 km, and mentioned marked changes in the height of reflection during a magnetic storm. Caton and Pierce (1952) observed that sferic waveforms arising from the southwest were different from those arriving from the southeast, which seemed to indicate that the earth's magnetic field influences the propagation. Their reflection heights averaged 86 km at night. Daytime waveforms were not suitable for analysis. Hepburn and Pierce (1953) deduced a daytime height of reflection of 65 km, and a nighttime height of 90 km, from observations of the "slow tail" of sferics.

In 1953 Budden attempted to fit the accumulated VLF observations of field strength versus distance into a model consisting of a sharply bounded, homogeneous ionosphere and a perfect earth as a waveguide, neglecting the earth's magnetic field. A computer was used to calculate waveguide modes and by trial to select parameters which fitted the observed field patterns. He found that the four least attenuated modes explained the measurements if a height of $69.1 \pm .05$ km was assumed, together with a collision frequency of 1 to 3×10^7/sec and an electron density of 135 to 400/cm^3 at the boundary. This was a daytime model.

In a later analysis, Wait (1957) showed that some of the modes (called negative order by Budden) had no physical significance. Using the corrected form of the mode expansion and a sharply bounded lower ionospheric model, Wait compared calculated field strength versus distance curves with the best available experimental data for daytime trans-Pacific paths. For frequencies in the range 16 to 25 kc the agreement was good for all distances if the collision frequency was taken as 10^7/sec and the electron density as 1,000/cm^3 at a reflection height of 70 km. At night the reflection height deduced was variable, but it seemed to be near 90 km. Also, some calculations by Wait and Perry (1957) of ionospheric reflection coefficients for a sharply bounded, homogeneous medium, with the magnetic field included, show a change of polarization and of phase of the $_{//}R_{//}$ coefficient with range indicating a quasi-Brewster angle effect which may explain the existence of the changeover in reflection characteristics observed by the earlier workers at intermediate ranges.

In summary, a complicated stratified region is probably not needed to explain the VLF results, but if the medium is so complicated, there may be subtle VLF effects as yet unobserved. As an example of such refinements, Wait (1958) has found it expedient to introduce a two-layered model in order to predict the field strength versus distance characteristics at frequencies lower than the range 16 to 25 kc considered in the earlier paper (Wait, 1957).

2.4. Theoretical Work Applying to Propagation of VLF and LF

A lack of analytical solutions of the equations of wave propagation in inhomogeneous anisotropic media for the general case applying to the lower ionosphere has plagued all experimenters. However, approaches have been made in two ways; first by finding approximate solutions of the wave equations, and secondly by devising digital computer techniques for integrating numerically the wave function through the medium. In both cases models of the lower inonosphere must be proposed, tried, and modified until the results fit the observed characteristics. With reference to stratification several workers have argued that split echoes do not necessarily imply stratification, since their theoretical work predicted a "coupling echo" in addition to the regular echo from a single layer. See Nertney (1952), Gibbons and Nertney (1952), Nertney (1953), Davids and Parkinson (1955), and Parkinson (1955). However, Budden (1955) in a complete treatment and some trial solutions by the numerical integration method, did not mention the extra echo. If the phenomenon exists, and his matrix $\overline{R}$ were plotted against height, it should exhibit the contribution from the extra "coupling" echo in appropriate levels. Poeverlein (1958) presents an analytic method consisting of approximation by very thin slabs which, according to the author, reduces to values equivalent to Budden's matrix $\overline{R}$. Both Budden and Poeverlein demonstrate a "quasi-Brewster angle" phenomenon with changing angle of incidence. (See comments referring to Wait and Perry (1957) above.) In summarizing this and the previous section, it should be noted that none of the methods discussed have enabled unambiguous, direct deductions of values of electron density and collision frequency versus height to be made. These should be the fundamental objectives of all ionospheric research.

2.5. Experiments With VHF Ionospheric Propagation

Research in VHF ionospheric propagation soon revealed that the signals were in general being returned from below the *E*-layer. Bailey, Bateman, and Kirby (1955) listed daytime heights between 75 and 80 km and nighttime heights of 85 to 90 km. The lowest height observed in some pulse tests was 59 km, and often several "strata" were simultaneously active. Pineo (1956) reported heights of 70 km by day and 86 at night, with frequent observation of two daytime reflection heights of 73 and 89 km simultaneously. It is quite apparent that these discrete heights at VHF cannot be explained by magneto-ionic splitting, and although there may be other difficulties of interpretation, these observations refute the idea that stratification is physically impossible in the region.

2.6. Lower Ionosphere Pulse Techniques at MF, HF, and VHF

Dieminger (1952 and 1954) called attention to the fact that low-level echoes were being observed in Germany by a conventional ionosonde. He observed the following characteristics:

(*a*) The echoes could be seen from 1.6 to 4.0 Mc at heights of 75 to 90 km in the daytime.
(*b*) There was a diurnal variation of height with a minimum at noon.
(*c*) The low levels merged with a 95 km level at night.
(*d*) The heights of all echoes were independent of frequency.
(*e*) There was a correlation of the absorption of higher level echoes with the height of the lower levels.
(*f*) Low level echoes were most frequent in winter, and occurred in groups of days.

In Dieminger's opinion the echoes were produced at a sharp boundary, but the echo structure was patchy, indicating an inhomogeneous medium. Meteor dust was proposed as the cause of a permanent nighttime echo at 95 km. He also concluded that the results were strongly influenced by meteorological phenomena, and that synoptic data over the world might be necessary to understand them. At about the same time, Gnanalingam and Weekes (1952) reported heights of 75 to 80 km on winter days of high absorption, but on other winter days heights of 90, 96, and 112 km were common. They used a sensitive frequency modulation technique to detect the weak signals. Later Gnanalingam and Weekes (1955) speculated on the nature of the region responsible for their observations. They argued that the echoes were not caused by gradients, because gradients would be destroyed by diffusion above 80 km, and gradients alone would not explain the absorption effects impressed on reflections from higher regions. They required a separate layer in the *D*-region which became totally reflecting at times at their frequencies (1.4 Mc). A histogram of all observed heights near noon during January and February 1952 and December 1953, showed no preferred values. The lowest height was 74 km.

Some lower ionospheric information has also resulted as a byproduct of auroral radar studies. Forsyth, Currie, and Vawter (1953) reported backscatter on 56 Mc which they deduced came from a level of 85 km. Similarly McKinley and Millman (1953) occasionally received vertically incident echoes on their 33-mc meteor radar at an approximate height of 80 km, night and day.

The first pulse experiment especially designed to study the lower ionosphere at vertical incidence at a fixed frequency above all critical frequencies of the region was described by Gardner and Pawsey (1953). They used a fixed frequency of 2.28 Mc and an overall sensitivity 30 to 40 db greater than Dieminger. Daytime echoes were regularly found at heights of 60 km and up. Field strengths of the different components differed by factors of up to 10⁴ and no echoes were ever observed below 55 km. Echoing layers seemed to come and go at fixed heights. Their heights were classified in groups at 70 km, 90 km and normal *E*-layer levels. Histograms of heights were also shown, arranged by the hour, and by the day. The lowest height showed a pronounced diurnal variation. The echoes returned from levels below 76 km were composed of a larger component of extraordinary polarization than of ordinary, suggesting that these were, indeed, partial reflections from less-than-critical density clouds of electrons in the region. In pursuing the analysis of the phenomenon Gardner and Pawsey show the theoretical basis for, and use of, a method of deriving a complete profile of electron density versus height in the region. Derived profiles are given for noon on two days in May. One shows maxima at 64 and 73 km. The other has maxima at 68 and 73 km. The authors conclude that most absorption of HF waves is above the 70-km region and that this region, although at a definite level having separate existence, is under some kind of solar control. They think that VLF propagation characteristics can be explained, using their results.

It should be mentioned that Appleton and Piggott (1954), being interested in the absorption of HF waves, investigated Dieminger's observations of low level echoes in winter by studying other regular ionograms. Again they found that occurrence on winter days of high absorption is directly correlated with the occurrence of sporadic reflecting strata below the level of the *E*-layer.

Gregory has used a vertical pulse equipment at 1.75 Mc of considerably greater sensitivity than that of Gardner and Pawsey, together with an antenna of 12 db gain directed vertically. His first paper (1956) gives results which not only confirm the stratification observed by Gardner and Pawsey, and Dieminger, but also show much additional detail as a result of the greater sensitivity. This enables him to deduce or substantiate the following:

(*a*) The existence of a modification of the *E*-region at 95 km, capable at times of total reflection, and showing some measure of solar control.
(*b*) The continuous existence, night and day, of a second region, having a lower boundary near 85 km.
(*c*) The existence of a complex series of partial reflections, during daytime only, extending down to 53 km, with a marked increase in strength and rate of occurrence at the lowest heights during winter months.
(*d*) The complete absence of reflections from heights below approximately 50 km.
(*e*) A relation, on winter days of high absorption only, between *E*-region echo strength and partial reflections from the heights 66 to 85 km.
(*f*) The findings of Dieminger that the low level echoes had a tendency to occur in groups of consecutive days, perhaps accounting for the winter absorption anomaly.

Observations of discrete heights at 95, 85, 75, 67, and 53 km seem common in Gregory's work.

Gregory (1957) published a comparison of heights observed at 1.75 mc with those reported by Pineo (1956) at VHF oblique incidence, pointing out the remarkable similarities between the two with respect to heights and diurnal variations. The suggestion was made that the two approaches (MF and VHF) should be regarded as complementary ways of investigating the nature of the region. A new low height commonly occurring on winter nights was found by Gregory (1958a) to be near 75 km, as contrasted with the 85-km limit observed by the earlier workers.

Gregory (1959) states that equipment for recording *D*-region echoes has recently been installed at Scott Base, 78°S, in Antarctica, and that reflection heights similar to those in temperate latitudes are being recorded. Initial results show that the all-daytime summer heights lie between 62 and 90 km, with the 90-km height present continuously. The fading rates are faster than at temperate latitudes.

Landmark (1958), in observations at 2.3 Mc of three polar blackouts, found that 65- to 70-km echoes appeared, but only when the main echo was increasing in amplitude (recovery period). The statistical significance of three cases may be somewhat doubtful, but polar blackouts are another phenomenon of great interest to which too little attention has been paid.

Bowles (1958 and private communication) has accomplished what amounts to a *tour de force* in the field of vertical incidence observations using pulses. With a transmitter capable of 4 to 6 megawatts peak pulse power, together with a 1024 element directive antenna, he has succeeded in obtaining reflections at 41 Mc which extend from the lower ionosphere continuously through to above the *F*-layer during the daytime. In particular, echoes, attributed to turbulence, are observed continuously at 60 to 90 km height. At any time several different strata are observed, but their continuity in time has not been established.

3. Rocket Observations of Electron Density—Height Profiles

The use of rockets over the past six years has introduced a powerful new tool in the study of the lower ionosphere. Much which in the past had to be inferred by remote radio and optical probing can now be measured directly. In this section only those rocket flights giving information on the electron density-height distribution are discussed. Other flights designed to give information on wind movements or on photoemissive processes are included in following sections on these subjects.

While rocket flights can give results on electron concentrations in the lower ionosphere, they must be treated with caution, because: (*a*) Experimental techniques are complex, and in some cases have known possibilities of error; (*b*) instrumentation has been adjusted for electron concentrations of the order found in the *E*-region. Consequently strata appreciably below *E*, of roughly 100-fold lower concentration, have hardly been recorded at all; and (*c*) the total observing time for all rocket flights is still extremely small, and hence a knowledge of the basic structure could still to some extent be obscured by local abnormalities.

Nevertheless, sufficient results are now available to show a fairly consistent pattern above about 88 km, with small concentration peaks and ledges superimposed on a steadily rising electron concentration with height. The rocket usually traverses a different part of the ionosphere on the upward and the downward portions of the flight, and either this space separation effect, or a fading or turbulent effect arising in the few seconds between the two portions of the flight, is sufficient in some cases to remove, or to alter the height of minor concentration peaks.

In 1953 Lien and coworkers reported four profiles, which differ in appearance, but three show a concentration peak at about 97 km, and two show peaks at 111 and 122 km. Observations were continued by Seddon, Pickar, and Jackson (1954), with peaks appearing at 101, 113, and 128 km. During the descent the 101 peak did not reappear. Jackson and Seddon, working with good equipment in 1958, find a strong peak at 101 km due to a sporadic-*E* layer, not more than 1 km thick. Again Pfister and Ulwick (1958) show peaks for one flight at 106, 111, 117, and 128 km, and comment on intense irregularities.

Jackson, in 1954, obtained daytime free electrons as low as 73 km. This was supported by Seddon, in 1958, who obtained a measurable electron density at 74 km, and also observed an abrupt ledge at 86.5 km.

4. Influence of Winds and Turbulence

4.1. Experimental Evidence

(*a*) *Radio echoes.* The ionosphere exhibits the characteristics of movement at all levels. As an example, sporadic-*E* clouds can be traced as they move bodily from one geographical location to another. The paper by Thomas and Burke (1956) gives an account of some of the deduced characteristics of these clouds. Their heights are above 100 km, their thickness is probably not greater than 1 km, and the horizontal extent is of the order of 10 km. It may be that similar structures can occur in the lower ionosphere.

Experimental work in studying motion of the lower ionosphere has invariably produced evidence that the apparent velocities change with height, sometimes very abruptly. Briggs (1951) found irregularities to be localized in height, using a frequency sweep technique which produced an interference pattern independent of ionospheric motion. By using a low frequency (150 kc) and supplementing the usual amplitude pattern with a phase height measurement, Jones, Millman, and Nertney (1953) reported that their fading was produced at heights

of 74 to 77 km by day, and 83 to 100 km at night. Another report of 150-kc observations was by Banerji (1955), who found day heights for irregularities to be lower than at night (78 versus 87 km). Bowhill (1956) studied fading phenomena observed from 16 to 2,400 kc, and suggested that a model containing small irregularities above the 70-km level and large ones below this level might explain the fact that a mixture of two different types of fading was observed. Later, Bowhill (1957) in a study of nighttime reflection levels of 75- and 150-kc pulses, came to a similar conclusion, but decided that the more rapid fading was due to irregularities above 100 km. Parkinson (1956), dealing especially with the nighttime 90-km echo, also found that it had two different characteristic fading frequencies, preseumably impressed upon the wave at two different levels. I. L. Jones (1958) reports that on two nearby frequencies, 2.0 and 2.5 Mc, the wind velocities were significantly different, perhaps because the fading was imposed by scattering near their different reflection levels. Also, the directions of the winds observed at the two frequencies rotated diurnally, the rotation at the greater height (higher observing frequency) leading that of the lower height. The phase lead amounted to 7 or 8 degrees per kilometer of height, thus showing wind shear.

(*b*) *The movement of meteor trails.* Again an appreciable body of literature has grown up in the past decade on meteor trail movements. Radar echoes from the trails of large meteors provide a powerful and direct method of observing movements in the 80- to 110-km region. Both steady movements, varying in direction with height and time, and turbulence are observed. Typical results are given by Greenhow (1952 and 1954) and Greenhow and Neufeld (1955 and 1956). They find that turbulent winds always occur between 80 and 100 km. There is a gradient of about 2.7 m/sec/km, but large irregularities in speed and direction are observed in a few minutes—occasionally amounting to complete reversals. Root mean square values of turbulence of about 30 m/sec are almost always observed. In their second paper they find that the height gradient of prevailing velocity was uniform at 0.5 m/sec/km ht.

A full statistical study of the accuracy of meteor wind measurements has been made by Manning, Peterson, and Villard (1954), who find strong evidence of layered horizontal winds in the 80- to 100-km region, with some superimposed irregular turbulent motion. Liller and Whipple (1954), photographing the movement of meteor trails, support these results by finding circulatory or turbulent winds moving in opposite horizontal directions separated by short vertical distances. Observed vertical wind shears averaging a greater maximum than 50 m/sec/km are not particularly concentrated at any height.

The literature contains many photographs and reports of long persistent meteor trails, showing in every case large amounts of relative movement. An excellent example is afforded by the artificial sodium trail created during a rocket flight (Edwards, Bedinger, and Manring, 1955). A photograph of the trail shows extensive shear distortion, and a complete reversal of wind direction was observed between the 85 and 110 km levels. These observations constitute convincing evidence that the motions are indeed true winds in the lower ionosphere. Other evidence is the motion of noctilucent clouds and airglow cells (Ludlam, 1957, and Roach, et al., 1958) discussed below.

4.2. Theoretical Aspects of Turbulence

The implications of the existence of turbulence and/or winds in the lower ionosphere have been studied because of the large influence they have on radio propagation. Herlofson (1951) pointed out that plasma resonances similar to those predicted for meteor trails could be expected for any small ionospheric irregularity. As a direct consequence, single-ray fading from this cause should be more important near the critical reflection level, and one should be able to identify a change of reflection level with a consequent change in fading characteristics by changing the probing frequency. As an extension, partial reflections from irregularities would be expected to show rapidly varying fading characteristics near the plasma resonances of the irregularities as the frequency is changed. (See the discussion of I. L. Jones (1958) above.) Polarization of the reflected wave could perhaps be examined to detect elongated field-alined irregularities.

In a paper by Maxwell (1954) turbulence in the *F*-region is primarily dealt with, but the author makes the statement that "the *E*-region will invariably be turbulent" with a microscale (small eddy size) of 10 m.

Gallet in two papers (1951 and 1955) develops a theory using the physical variables in a turbulent region to explain some of the radio reflective properties of the *E*-region, first as an explanation for the sporadic-*E* reflections, and then for some of the lower ionospheric phenomena. In a model region, vertical transport of eddies or "blobs" of air changes their pressure, density, and temperature. The contrast in electron density presented by a blob against the background of its new location is a function of two sets of circumstances. First, how different its electron density was before it was moved with respect to the surrounding electron density of the new location, and secondly, how different a volume appears to be to the exploring radio wave due only to the changes of pressure, density, and temperature with height. The first situation is obviously of maximum effect near a sharp electron density gradient or, in the case of a single layer, above and below the maximum. The second situation requires no gradient of the average electron density, and therefore requires no layer; and the position of maximum contrast would be entirely a function of the localizations, if any, of the turbulence characteristics. In application, these two mechanisms could give maxima at different levels, recalling the question of the interpretation of split echoes from the lower ionosphere.

The idea of turbulence due to ionospheric wind motions has become a very useful one in attempting explain many ionospheric phenomena that have been known, but seldom studied. Booker (1956) has listed many phenomena that he considers can be explained by turbulence in the ionosphere, including fading of single rays, the distortion of meteor trails, *D*-region echoes and VHF forward scatter. Booker conjectures that the latter two are caused by the same irregularities. (See Gregory (1957) and Bowles (1958).)

As an example of an interesting application of the gradient mixing type of turbulence theory (see Gallet above), Wheelon (1957) has tried a model of the lower ionosphere which might explain the VHF forward scatter observations, particularly the diurnal variations, as described by Bailey, et al. (1955) and Pineo (1956). The model starts with a very steep gradient of ionization at 70 km of 2,300 electrons/cm^3/km to provide the observed daytime field strength, and includes another rise between 75 and 88 km to account for the nighttime residual signal after the 70-km daytime component has decayed.

The existence of winds and turbulence in the lower ionosphere may provide a coupling link between that region and the next lower region. As an example of the type of meteorological work which may eventually establish a link between the two, the paper by Fleagle (1958) is of interest.

5. Survey of Meteor Height Results

Extensive studies of meteors entering the earth's atmosphere have been made by visual, photographic and radio echo methods, particularly in the last decade. This work gives information relevant to the present inquiry in the fields of meteor ionization height, and the influence of fine meteoric dust.

Many papers report studies of the height of occurrence of the light and the ionization from meteor trails. The mean height of maximum ionization of visual size meteors is close to 95 km (Hey and Stewart, 1947; Millman and McKinley, 1949). Analyses of both visual observations, given by Porter (1944), Prentice (1948). and others, for the beginning and end heights; and of radar echoes given by Evans (1954) and others for the height of maximum ionization, show that these are clear functions of meteor-velocity. The height of maximum ionization for the smallest detectable meteors falls from 105 to 85 km as the incident meteor velocity decreases from 60 to 20 km/sec.

The effect of initial meteor size on the height of maximum ionization is much less marked than the velocity effect. Prentice, recording visual results, finds that large meteors extend over a greater height range than small meteors, but Millman (1957), observing radar echoes in Canada, finds no mean variation in height between meteor magnitudes of −6 and +4. From this Millman concludes that the atmosphere is the controlling factor in determining the height of maximum ionization. This result agrees with the work of McKinley and Millman (1949) who, on the basis of a considerable body of radar echo observations, conclude that the 90- to 100-km region has fine structure, or localized regions which prolong the existence of meteoric ionization at specific heights. As a result of further analysis, Millman (1957) reports the possible occurrence of peaks of ionization at 92, 98, and 106 km.

Herlofson (1948), however, deduced theoretically that the atmospheric pressure at the point of maximum luminosity of a meteor is directly proportional to its pre-atmospheric radius. Since pressure is a simple function of height, this work implies that smaller meteors create their maximum ionization at a greater height. This conclusion is supported theoretically by Kaiser (1953), who places the maximum ionization effect of small meteors at heights between 130 and 115 km.

There is obviously a discrepancy between the smooth transition to greater heights with decreasing meteor size predicted theoretically, and the height independence of size, and the stratification, observed by the Canadian workers. However, it can at least be concluded that practically all significant meteoric ionization produced by large to near micrometeoric size particles, appears well above the mesopause height of 82 km, and well clear of the ionospheric stratification observed by Gregory at still lower heights.

Micrometeorites have been shown by Whipple (1952) to be so small that they are capable of deceleration to zero velocity by the drag of the earth's atmosphere, without ablation occurring. Ablation was necessary in Herlofson's model for the production of ionization. Consequently micrometeorites were disregarded as a source of ionization until Dubin (1955) produced evidence that highly velocity-sensitive ionization could be produced without ablation. The number of such particles is great, and on extrapolation from the Herlofson and Kaiser theories such ionization, if produced, should occur at still greater heights than that produced by the smallest ablating meteors. (Dubin has assumed, without justification, that the micrometeorite effect occurs at the height of sporadic *E*.)

On this basis Gregory has assumed, along with others before him, such as Deb (1940) and Paton (1951), that meteoric dust residues, and presumably also micrometeorites, would drift down to and concentrate at the temperature inversion level at about 82 km. There, the mechanism of photoionization of dust particles by solar radiation, suggested by Eckersley in 1948, might give rise to further ionization during the daylight hours.

The subject, however, is now further complicated by the recent analysis of Öpik (1958), who concludes that micrometeorites reach directly to, and have their maximum temperature in the region 70 to 90 km. If Dubin's suggestion—by analogy with shock-tube experiments—that such neutral particles can indeed produce appreciable ionization by velocity alone, is accepted, then a new source of low-level ionization is apparent. Further experimental work

on the variation of the height of maximum meteor ionization with decreasing particle size is clearly required.

Evans (1955) in the northern hemisphere, and Weiss (1955) in the southern hemisphere, find no seasonal effect in the height of meteoric ionization.

6. Summary of Optical Results Giving Height Information

Observations of scattering of the sun's rays by the atmosphere at twilight give evidence of discontinuities in the lower ionosphere. Even at night the sky is not without light, but always has a faint glow. In the past 5 years the heights of emission of various spectral lines in the airglow have been determined with increasing certainty. On occasion noctilucent clouds are also visible at great heights. These optical phenomena have given, with great consistency, a height of 82 to 85 km.

6.1. Noctilucent Clouds

As early as 1934 Hoffmeister attempted to relate the occurrence of luminous bands in the night sky with ionospheric variations by assuming that both variations were due to sudden invasions of meteoric dust.

Paton (1951), reporting on true noctilucent clouds, considered that they were usually blue, indicating very small particle size, but that some portions had been known to turn white. This change was attributed to particle growth by condensation of water vapor on meteoric dust. A very comprehensive survey of all available information on noctilucent clouds was given by Ludlam in 1957. The spread in height was between 79 and 90 km, with the mean of 179 observations at 82.1 km. The particle size was in general too large to give a blue cloud by Rayleigh scattering. The temperature minimum at the inversion height, where the clouds are found, was considered to be too high to allow condensation of water vapor at the prevailing air pressure unless a very abnormal humidity condition was postulated. Once again, therefore, the material was considered to be dust. The specific height is attributed to solid material falling from above 82 km due to gravity, or rising from below the temperature inversion due to reradiation of solar energy to the surrounding air, with consequent upward movement of both air and particle. Noctilucent clouds have a considerable directed horizontal drift speed.

6.2. Scattered Sunlight

At dawn the sun progressively illuminates lower regions of the ionosphere, or the converse at dusk, while the earth remains in darkness. Two techniques have been used to observe the ionosphere at these times.

The first, giving surprising results, was carried out by Khvostikov and Sevcenko as long ago as 1936 in the Caucasus. Basically, their assumption was that a strongly ionized gas would affect the polarization of the diffused light, and that at any given moment at dusk the greater part of the diffused light proceeds from a relatively thin region. The polarization of the diffused light from the zenith was measured continuously, at 3,000 m above sea level, and discontinuities in the degree of polarization versus height observed. Due to unfavorable weather conditions, only seven curves were obtained, but these are just enough to indicate that similar heights were observed on different days. Five morning curves show heights of 95 and 135 km, and two evening curves indicate heights of 80, 115, and 150 km. (These results were obtained before the *E*-region was considered to be other than a single layer.)

The higher morning values were attributed to nighttime recombination. The results, admittedly, are no more than an indication of a method, but the technique does not appear to have been used since 1936. The method is quite simple, and in a suitable locality may well be worth further study, particularly as it can reveal small maxima of electron concentration above a larger maximum.

The second technique, due to Bigg in 1956, involves direct twilight scattering. He was able to show a small maximum in the scattered light at 81.1 ± 5 km, and assumes it is due to a meteoric dust layer.

6.3. Airglow

The night sky radiates a weak continuum, together with a number of emission bands. In addition certain lines, especially of sodium, show a more prominent emission at twilight. The airglow, excluding the sodium and calcium twilight emission, will first be considered.

Many early results on the height of the airglow have been rendered obsolete by the work of Bates and Dalgarno (1953), who show that most mechanisms can only apply below 100 km; together with a fuller recognition of the experimental difficulties in this work, and the necessity to eliminate all light of astronomical origin.

Most work has been done on the green OI radiation of wavelength 5577 A, in the airglow. In 1955, St. Amand, Pettit, Roach, and Williams were able, by triangulation, to place the height at 80 to 199 km. In the same year Roach and Meinel listed the spread in heights of several airglow components. Each of the components occupied a different height zone, and the green OI line was found between 62 and 104 km. Elsasser and Siedentopf, in 1956, place the green OI line at 90 ± 10 km, and Roach, Megill, Rees, and Marovich (1958) at about 100 km, agreeing reasonably with a rocket flight in 1956 (Berg et al.) where the peak luminosity was found at 90 to 95 km, and another flight in 1958 (Heppner and Meredith) where the radiation was observed only between 90 and 118 km, with a sharp lower boundary. Manring and Pettit (1956) find variations of the height with night and season, so that definite heights—using the accepted Van Rhyn method of reduction—are not possible at present.

Night NaD was found by Roach and Meinel between 108 and 129 km. This fairly wide spread is to some extent reinforced by Heppner and Meredith's rocket flight, where NaD was found primarily between 85 and 110 km, with a maximum at 93, but with a tail-off in distribution extending into the 130 to 150 km region.

In two papers Roach, Tandberg-Hansen, and Megill (1958 a, b) have reemphasized that the night airglow as seen by the green OI radiation is not uniform in its distribution, but is concentrated in cells or patches whose characteristic size is approximately 2,500 km in horizontal extent. During the period October–December 1956 the cells seemed to move, usually from north to south with a mean speed of about 100 m/sec. The authors, in the second paper, have decided that the apparent motion is due to actual transport of particles carried along by large-scale wind movements, and point out that the composite picture is generally similar to vortex cells in the troposphere. For a phenomenological description of the night airglow, including historical background, see Roach (1959).

6.4. Sodium and Calcium at Twilight

Hunten and Shepherd (1954) find a maximum sodium density at 85 ± 3 km, with an exponential decrease above and below this height. Hunten, in the following year, surveying two years' results on the height of maximum density of the sodium, found that the height remained constant within 1 or 2 km of 84.5 km, but that the intensity varies seasonally, and concluded from this that there may be transport of neutral sodium from polar regions in winter. Finally the work of Chamberlain, Hunten, and Mack (1958) supports the view that twilight emission is a resonance scattering effect, rather than a dissociation brought about by ultraviolet light. It is possible that the sodium travels up through the whole atmosphere from sea level. A latitude effect also exists. Correlation of abundance changes with the passage of meteor streams may yet be observed, but until it is there is no necessity to invoke a meteor origin for the sodium.

A. V. Jones (1956 and 1958), however, has observed with certainty twilight emission from ionized calcium, possibly correlating with meteor activity, particularly during the August Perseid shower. There is some evidence that the emission is concentrated in the 100- to 120-km region. This height happens to agree with the value calculated by Kaiser for maximum integrated meteor ionization, and the height of 100 to 110 km for the maximum of the night glow continuum determined by a rocket observation (Heppner and Meredith).

No author appears as yet to have given an explanation for the appreciably greater height of night sodium than the twilight sodium glow. As both sodium and calcium are constituents of meteors, identification of intensity variations in the night sodium—or even possibly in the twilight glow at lower height—may yet be found to be a function of meteor activity, similar to the calcium activity.

Heavy atomic ions, such as calcium, may also be introduced into the atmosphere following solar flares (Bailey 1957), and this could both complicate the variability of the twilight glow, and provide a temporary source of extra ionization.

7. Photoemissive Processes

The literature abounds with explanations of the production of free electrons by the interaction of selected solar radiations with specific upper atmosphere constituents. In conjunction with other theories explaining various recombination processes, these lead to the desired equilibrium concentrations of free electrons at specific heights. The field of study is an extremely difficult one, requiring as it does a combination of detailed knowledge of atomic processes, of the flux of incoming solar radiation, and of the exact components and physical conditions in the upper atmosphere. Many of these explanations have subsequently been disproved, and a new discovery, such as the existence of appreciable quantities of solar X-radiation, entirely altered the position. In view of this prevailing uncertainty, only brief mention will be made here of some recent views of electron production and decay in the lower ionosphere.

Chapman's theoretical treatment (1931) of the ionizing effect of monochromatic radiations of known intensity on particular atom species in an idealized atmosphere is still basic to most theories.

7.1. *E*-Region

Ionization at a true height of about 130 km was for some years considered to be due to photoionization of atomic oxygen by ultraviolet radiation ($\lambda > 910$ A, Nicolet 1952). By 1954, Bates regarded this process as accounting for the E_2 layer only, and reintroduced the photoionization of molecular oxygen as the main contributor to the E_1 layer, at the same time recognizing that X-rays could be of significance. New data for the absorption of radiation in the 850 to 1100 A band by O_2 and N_2 led Watanabe, Marmo, and Pressman (1955) to conclude that Lyman β could ionize O_2. Bates was thus supported, and the ultraviolet ionization of O_2 was not entirely rejected in favour of the newer X-ray theory of E_1 origin. By 1956, however, Bates had concluded that X-rays of about 40 A were *mainly* responsible for the *E*-region, and the theories of ultraviolet ionization of atomic oxygen, together with the longer wavelength X-ray reactions, were transferred to the F_1 and F_2 regions.

7.2. *D*-Region

A somewhat greater uniformity prevails in theories of the origin of the *D*-region. In 1945, Nicolet suggested that the causitive agent was the photoionization of nitric oxide, according to $NO + h\nu$ ($> 9.5eV$) $\rightarrow NO^+ + e$ by Lyman α radiation (λ 1216 A)

which is able to penetrate the atmosphere to a height of 70 km.

A considerably increased value for the ionization cross section of NO, obtained by Watanabe, Marmo, and Inn (1953), reinforced Nicolet's original suggestion. Nicolet (1954) arrived at the specific conclusion that the Lyman α photoionization of a concentration of NO of up to 10^8 molecules/cm^3 caused the i -region at a height of less than 80 km.

A model of the D-region using this reaction was next proposed by Mitra (1954), with an assumed NO distribution, and a maximum NO concentration approaching 2×10^{12} molecules/cm^3 at 85 to 95 km. A further assumption of a recombination model was necessary to give an electron distribution in agreement with that deduced from certain radio observations. Direct rocket measurements (discussed by Bates, 1956) give a relative volume abundance of NO of $\leq 2\times10^{-4}$. If the tota atmospheric concentration of all atom species is taken as 8×10^{13}/cm^3 at 90 km, the NO concentration would be $<4\times10^9$/cm.[3] The corresponding figures at 80 km are of the order 4.8×10^{14}/cm^3 and hence $< 2.4\times10^{10}$/cm^3 for NO. Thus the concentration proposed by Mitra is clearly greater than observed, but Nicolet's figures remain feasible. Rocket observations (Friedman and Chubb 1955) confirm the presence of Lyman α in the D-region, with solar X-rays contributing to the upper part.

Finally, Nicolet (1954) has introduced atmospheric motions, causing fluctuations in NO concentration, to account for the variations in the D-region ionization

7.3. Sudden Ionospheric Disturbances

Various radio methods have long shown with clarity a sudden increase in ionization at very low heights, down to 64 km, during a solar flare. As an example, Gregory (1958 b) presented an analysis of the effect of SID's on 1.75 mc reflections. In 7 cases out of 18 all echoes vanished during the disturbance; in 11 cases one height only (56 to 65 km) remained, and in 9 of these 11 this remaining height had not been present prior to the disturbance. He concludes that the occasional disappearance of all echoes indicates that absorption without reflection has been produced below 56 km.

It remained for rocket flights to confirm the increase in Lyman α and X-rays during such times as postulated by Bates and Massey (1951) and Piddington (1951). nhanced Lyman α intensity and X-radiation $\sim$ 2 A have now been observed down to the 60 to 70 km region in a typical flare (Friedman and Chubb, 1955; Chubb et al. 1957). These observations lend support to the importance of both Lyman α and X-rays as normal ionizing agents in the lower ionosphere.

7.4. Recombination

The equilibrium electron concentration is also a function of the recombination rate. In this field theories do not as yet provide conclusive explanations. Direct (radiative) recombination, expressed in the simple equation $(dN/(dt)=q-\alpha N^2$ where N is the electron concentration, and q and α the effective rates of electron production and recombination respectively, does not agree with the observed rates of electron density change. The exponent 2, commonly used, is open to question (Rawer, 1956). Negative ions may play an important part below 80 km (Mitra and Jones, 1954), and much patient unraveling will be required to give a picture consistent with observed electron concentration values in the lower ionosphere.

8. Discussion of Survey Papers Dealing With the Lower Ionosphere [3]

1. Bracewell et al. (1951), **The Ionospheric Propagation of Low and Very Low Frequency Radio Waves Over Distances Less than 1,000 km.** This paper contains 54 references. Observed reflection heights are tabulated. One paragraph (5.2) discusses the evidence for multiple strata observed at 70 and 113 kc

2. Mitra (1951), **The D Layer of the Ionosphere.** In part I the author gives a résumé of the present state of knowledge concerning the D-layer, and in Part II speculates on the origin and structure of the region.

3. Kitchen, Pressey, and Tremellen (1953), **A Review of Present Knowledge of the Ionospheric Propagation of Very Low, Low, and Medium Frequency Waves.** Some sferics work is included, and recommendations are made for further detailed research.

4. Briggs and Spencer (1954), **Horizontal Movements in the Ionosphere.** A survey of results obtained by radio methods. The authors decide that there are both irregular and systematic (diurnal and seasonal) horizontal movements at all ionospheric levels, and that there is a tendency for velocity to increase with height. Also, the velocities increase with magnetic disturbance, especially at the higher heights. The paper ends with a summary of non-radio methods.

5. Weekes (1954), **The Physical State of the Upper Atmosphere.** The author starts with a general discussion of the problems and then discusses the following specific fields:

Composition, temperature and pressure of the atmosphere at different levels, results of rocket measurement of those quantities, measurements of wind and turbulence, and transport of ionization in the upper atmosphere. Many references are given.

6. Waynick (1955), **The Lowest Ionosphere.** The main properties of the region are surveyed, with the conclusion that most of the experimental evidence is in favor of at least one D-layer below E-region, with possible additional stratification of one form or another. A collection of curves from different sources is given showing the variation of recombination coefficient with height.

[3] Material in the following papers has been included, where applicable, in the general text.

7. Murgatroyd (1957), **Winds and Temperatures between 20 km and 100 km—a Review.** The author states that the measurement of winds is a more reliable way of calculating temperatures in the upper atmosphere than vice versa. He proposes a standard atmosphere to 100 km based almost entirely on wind measurements, which is applicable between 30° and 60° latitudes. Before much progress in the study of turbulence can be made, a knowledge of the wind structure in considerably greater detail is necessary. In addition, measurement of atmospheric composition, including ozone and water vapor at all possible levels, is necessary to allow studies of the radiation of heat energy.

9. Tables of Observed Heights

In an endeavor to determine whether any pattern of heights exists, all the relevant experimental results so far discussed have been collected in tables 1 and 2, with the exception of lowered heights recorded during solar flares.

The results given in table 2 reveal two facts. First, while a definite height may be obtained for a specific process, different classes of phenomena show their maxima at different heights. Secondly, some weak grouping of events in height ranges may be discerned. During the daytime three, or possibly four height groups contain the major proportion of the results. These are the belts 67 to 75, 80 to 90, 95 to 97, and possibly 110 to 113 km. At night only the two center groups remain, namely 80 to 87, and a somewhat wider 92 to 98.

The only nighttime results below 80 km are Rivault's sferics observations, giving a mean height of 75 km; and frequent vertical pulse echoes in winter down to 75 km obtained by Gregory. Rivault reports a lowering to 70 km during magnetic storms, and it may well be that Gregory's low height night results, including especially one observation at 67 km, are due to a similar cause.

10. Characteristics of an Acceptable Model of the Lower Ionosphere

As an outcome of the experimental data assembled above, it becomes possible to specify with reasonable certainty a number of conditions which must be met by any model purporting to explain the lower ionospheric radio observations. These characteristics are as follows:

(*a*) Probably the most consistent feature is the presence, night and day, of a reflection height of about 85 km at temperate latitudes. This height is roughly the same for all vertical incidence measurements from very low, through medium, to very high frequencies. The height appears to be associated with the temperature minimum (mesopause), and is the same height at which sharp maxima are found for the twilight NaD radiation, the occurrence of noctilucent clouds, and the twilight scattering of sunlight from dust.

(*b*) Notwithstanding (*a*), the effective reflection height calculated from *oblique* incidence results appears to fall considerably with increasing range at very low to medium frequencies.

(*c*) Reflections at vertical incidence for medium frequencies and above are usually weak and hardly ever blanketing.

(*d*) On a sweep-frequency record, retardation cusps are not apparent on progressing from one reflection level to the next one above.

(*e*) Reflections from the region should not be polarized according to either magnetoionic component alone, but should have polarization corsponding to a mixture of the two.

(*f*) Several reflection heights usually occur simultaneously.

(*g*) The individual heights below the 85-km region can remain constant for periods of many hours, but may be different on different days. Observation of the lowest heights in the region requires a very large system sensitivity.

(*h*) Heights below 80 km are usually observed in the daytime only, and most often in the winter. Winter days of unusually low recorded heights occur in groups of several consecutive days.

(*i*) All reflections exhibit fading. The fading rate should also increase with increasing magnetic disturbance.

(*j*) During some SID's a new reflection height, lower than those previously seen, should be observed at medium frequencies, together with a weakening or complete disappearance of the signals from greater heights.

11. Model A

The observations showing the tendency for wind shears to exist in the lower ionosphere suggest that the discrete heights observed by radio reflection may be explained entirely by scatter from the turbulence thus produced. As discussed by Gallet (1951 and 1955) and Wheelon (1957), turbulent regions may become visible to radio waves when a gradient of ambient ionization is also present. Therefore the following model is proposed:

(*a*) The air is especially turbulent at certain altitudes. The altitudes are not specified, but there must be several and they must have the diurnal and seasonal variations compatible with the experimental evidence.

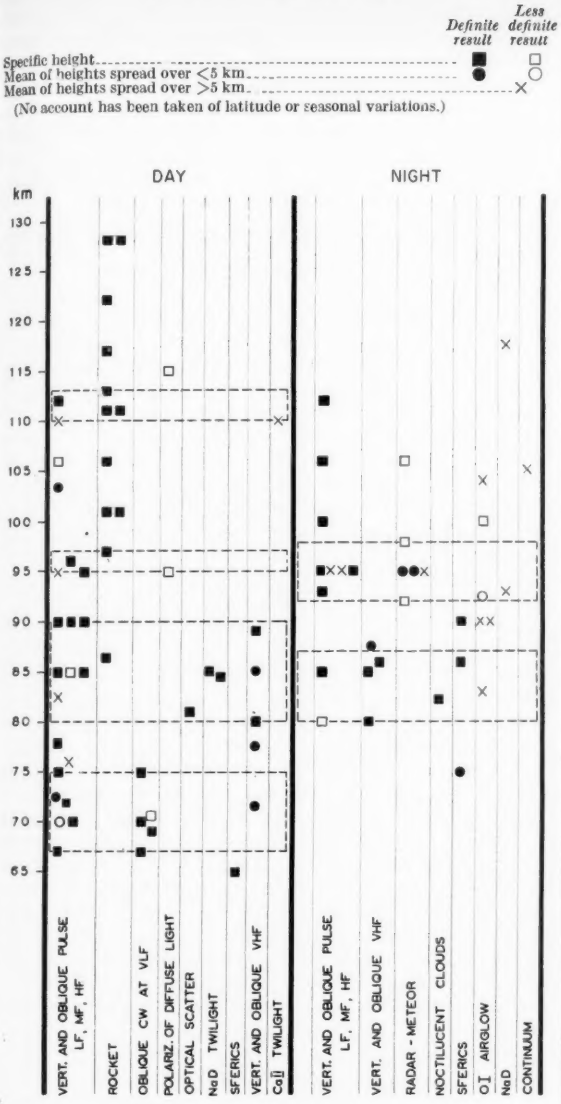
(*b*) A rather simple profile of ambient electron density versus altitude is assumed to occupy the whole region, having diurnal and seasonal variations due to radiation from the sun, and also showing occasional distortion due to solar eruptions. A possible arrangement is shown diagrammatically in figure 1.

The shaded areas are intended to represent the height, thickness, and (by their length) intensity of the turbulent regions. The regions will be most readily "visible" in the radio sense when there is a sufficient gradient of the ambient electron density—as given by the slope of the profiles—in these regions.

TABLE 1. *Observed heights*

Method	Authors	Year	Day	Night	Comments
			Km	*Km*	
Vert. pulse HF	Appleton and Piddington	1938	Broad peak at 110, nothing below 80.		
Vert. pulse HF sweep	Appleton, Naismith and Ingram.	1939	Partial at 90–100		Unrelated to Es.
Vert. pulse 100 kc	Helliwell	1949		84–106	Complicated split echoes.
Vert. pulse 100, 325 kc	Helliwell, Mailinkrodt and Kruse.	1951		90–130	Do.
Vert. pulse 100, 310, 2,400 kc	Helliwell	1952		93, 100, 106, 112	
Vert. pulse 150 kc	Lindquist	1953	85	95	
Vert. pulse LF sweep	Watts and Brown	1954	Layers between 70 and 100	90–100	
Vert. and oblique pulse MF	Naismith and Bramley	1951	70–75 equiv. vert. common at long range.		Day or night not specified. 85 occasionally, 106, 120–130.
Oblique pulse 300 kc	Watts	1952	~70	>80	
Vert. pulse inter-action 2,500 kc	Fejer	1955	72, 78		
Vert. pulse 1.4–4.0 Mc	Dieminger	1952	75–90 (winter)	95	More low level echoes in winter.
Vert. pulse 1.0, 1.4 Mc	Gnanalingam and Weekes	1952	90, 96, 112 (winter) 102–105 (summer)		75–80 on winter days of high absorption.
Vert. pulse 2.3 Mc	Gardner and Pawsey	1953	70–90 and *E*		Lowest height 55.
Vert. pulse 1.8 Mc	Gregory	1956	53, 67, 75 frequently, 85, 95 usually.	85	Lowest height 53. Low heights more common in winter. 85 always present.
Vert. pulse 1.8 Mc	Gregory	1958		Sometimes observed down to 75.	
Vert. pulse 2.3 Mc	Gregory	1959	62–90, 90 continuously		
CW oblique VLF	Weekes	1950	67, 75		Short range.
Do	Bain and Bracewell	1952	76		
Do	Budden	1953	69.1		
Do	Wait	1957	70		
Sferics	Rivault	1950		75 mean	70 or lower during magnetic storms.
Do	Caton and Pierce	1952		86	
Do	Hepburn and Pierce	1953	65	90	
Oblique pulse VHF	Forsyth, Currie and Vawter	1953		85	
Do	McKinley and Millman	1953	~80	~80	
Oblique CW and some pulse VHF.	Bailey, Bateman and Kirby	1955	75–80	85–90	Lowest day height 59, at several strata simultaneously.
Oblique pulse VHF	Pineo	1956	70–73, 89	86	
Vert. pulse 41 Mc	Bowles	1958	~85		
Rocket transmitter	Lien et al.	1953	97, 111, 122		
Do	Seddon, Pickar and Jackson	1954	101, 113, 128		
Do	Jackson and Seddon	1958	101		=Es.
Do	Pfister and Ulwick	1958	106, 111, 117, 128		
Do	Seddon	1958	86.5		
Radar meteor echoes	Hey and Stewart	1947		95 mean	Day and night not usually distinguished. Majority of results night-observed.
Do	Millman and McKinley	1949		95 mean	
Do	McKinley and Millman	1949		90–100 fine structure	
Do	Millman	1957		92, 98, 106	
Noctilucent clouds	Ludlam	1957		79–90, peak at 82.1	
Polarization of diffused light	Khvostikov and Sevcenko	1936	95, 135 at dawn, 80, 115, 150 at dusk.		
Scattered light	Bigg	1956	81.1 ±5		
OI airglow (5577 A)	St. Amand et al.	1955		80–100	
Do	Roach and Meinel	1955		62–104	
Do	Elsasser and Siedentopf	1956		90 ±10	
Do	Berg et al.	1956		Peak intensity at 90–95	
Do	Heppner and Meredith	1958		90–118	Sharp lower boundary.
Do	Roach et al.	1958		~100	
Night continuum	Heppner and Meredith	1958		100–110	
NaD twilight	Hunten and Shepherd	1954	85 ±3 peak		Exponential decrease on either side of peak.
Do	Hunten	1955	84.5		
CaII twilight	Jones	1956	100–120		
NaD	Roach and Meinel	1955		108–129	
Do	Heppner and Meredith	1958		85–110, peak at 93	Tails off to 150.

TABLE 2. *Grouping of heights*



The gradient must be dynamically stable, or the turbulence will, in time, destroy it.

Three electron density profiles, A, B, and C, are shown. Using A, the results of a normal day might be expected—reflections from several levels, of increasing gradient, and hence of increasing echo strength, with height. Profile B contains no electrons at the lower levels, corresponding to the disappearance of the lowest echoes at night. Profile C has a bulge below 80 km and might correspond to conditions during an SID, when the lowest reflection levels are enhanced and the other reflections weakened by increased absorption and the changes in slope along the profile.

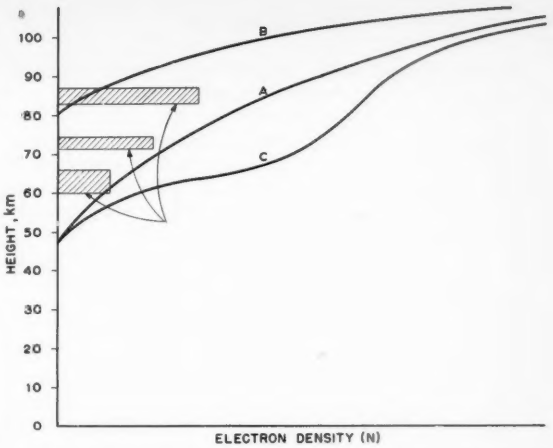


FIGURE 1. *Model of lower ionosphere showing strata of turbulence in relation to ambient electron density profiles.*

Arrows indicate regions of maximum turbulence.

It is suggested that this model is capable of fitting the experimental facts, at least qualitatively, as follows:

(*a*) Reflection heights at vertical incidence will be the same at all frequencies above the critical frequency for the ambient electron density at the reflection level. This is due entirely to the localization of the turbulent regions

(*b*) The lowest observable reflection height as determined by (*a*) will vary with the angle of incidence, since the lowest regions present more oblique aspects to the wave. Thus with increasing one-hop distances lower height echoes will gradual y appear.

(*c*) At vertical incidence reflection will be partial at frequencies above the critical frequency for the reflection level. Plasma resonances due to the irregularities (see Herlofson 1951) may be observed near the critical frequency.

(*d*) On a sweep frequency exploration, the levels will be penetrated without retardation, provided penetration occurs well above the critical frequency of the reflection level.

(*e*) The reflected wave will not be split into ordinary and extraordinary components because of the absence of selective retardation (see (*d*) above).

(*f*) All reflection will fade, showing little or no specular component. The rate of fading will be a function of the meteorology of the region of reflection.

(*g*) Lowest heights are a function of both the turbulence and the ambient electron profiles. Because of this dual character, echo levels and intensities may vary from day to day.

This model can thus explain the radio observations of reflection heights at discrete levels without requiring extremely sharp gradients of the average electron density. It does not preclude the appearance of discrete reflections from real strata of ionization.

12. Model B

An alternative model might be constructed on the assumption of multiple Chapman processes. In this connection a model of the *D*- and *E*-regions calculated by Houston (1958), after a quantitative reappraisal of the ionizing radiations and the atmospheric constituents involved, is of interest. The only significant cause of the *D*-region was found to be the action of Lyman α radiation on NO, and it is probably coincidental that the computed ionization maximum at solar zenith was 83 km, in agreement with other mesopause phenomena.

Solar energies, particularly in the 100- to 1000-A region, are still largely unknown and the possibility of significant new processes being found cannot yet be ruled out. An attempt to include this portion of the spectrum has been made by Kallman (1957), who deduces therefrom a total distribution curve commencing at 88 km and rising smoothly to 105 km. Again, as is shown by the night OI airglow radiation, processes not directly involving ionization occur at fairly specific heights. Such processes may play an indirect part in daytime ionization control. Further, the resultant distribution even for a single layer may be far from a Chapman type due to variations in absorption of the band of wavelengths producing a particular photochemical reaction. Recent rocket studies, however, have led to a clarification of processes, and new discoveries are becoming less likely, especially at low heights. No model as yet shows any sign of multiple strata below 80 km.

A strong argument in favor of a photoemissive origin of such layers is the production of ionization at 64 km during a solar flare. Increased X-radiation causes this effect, so a smaller amount of radiation at this or slightly higher heights, while producing a smaller effect, should still be detectable with high system sensitivities. Even model A requires that the SID effect be of photoemissive origin.

Again, any layer under solar control should exhibit a considerable height change during the day due to $f(\chi)$ in the Chapman equation, where χ is the solar zenith distance. This is clearly shown in Houston's model, where the height of the maximum *D*-ionization rises from 83 to 100 km as χ changes from vertical to horizontal incidence. Diurnal height changes have been found at 16 kc (Straker, 1955), and the height of the lowest partial echo at MF shows a similar change (Gardner and Pawsey). It is possible that diurnal variations in the height of a single low stratum have not yet been closely looked for, but one point is quite clear. The electron concentrations, of the order of 1,000/cm^3 or less, are far below the critical density required for classical MF or higher frequency reflections, and the phenomenon at those frequencies must be some form of scattering process.

Very sharp electron gradients can cause MF reflections, but the gradient values required are unlikely to be produced by solar effects alone. The simultaneous occurrence of several low strata, which have recently been observed in Germany, Australia, United States, New Zealand, and Antarctica, are surprisingly constant in height with time if the visual meteor trail turbulence observations of Liller and Whipple are accepted. The opposite case for stratified turbulence, however, appears from various radio echo experiments (sec. 4.1(a)). Consequently no firm conclusion can as yet be reached on the significance of photoemissive processes in the very low ionosphere.

13. Conclusions and Recommendations for Further Work

The authors conclude that the lower ionosphere exhibits the property of stratification as evidenced by all the available experimental results at medium and higher radiofrequencies. The fact that the properties of the region affecting low and very low frequencies may be explained by simpler structures does not appear to preclude a more complex structure, since the summation effects of long radio wavelengths would be expected to conceal details having physical dimensions much smaller than one wavelength. It is suggested, however, that rather subtle anomalies of long wave propagation may exist which will require more detailed knowledge of the fine structure of the region for their interpretation. In this respect, observations at a single low frequency have probably reached the stage of diminishing returns, and finer probing, such as is given by high power narrow pulsed HF systems will be necessary to reveal further details.

Evidence is strong for the accumulation of dust, and of sodium and probably calcium atoms at the temperature minimum at about 85 km in temperate latitudes. This accumulation may or may not be the cause of either an ionization maximum or a steep gradient at this height, but at least a partial reflection at MF is always obtained, provided the sensitivity is high enough. Beyond this fact it can only be said that strata exist over the region 54 to 130 km, without any obvious constancy of height or of height separation. There is no series of preferred heights.

Any further physical explanation of the fine structure can only be guessed at in the absence of extensive knowledge of its characteristics. It seems evident, out of all the material consulted, that only two or three experiments have given unambiguous information on the basic ionospheric quantities in the region. Acceptance of the evidence for stratification automatically casts doubt upon the meaning of the radio height results. The results from most pulse-type radio probing techniques in this region may be explained by any or all of the following: electron concentration maxima, electron gradients, or zones of turbulence.

Direct probing by rockets is one method which does provide the required electron profiles. If possible, it would be highly desirable to carry out a series of electron profile observations capable of recording the weak electron concentrations in the lower ionosphere, at a fixed site. If this is not

possible, rockets could at least be used in conjunction with other cheaper experiments in an attempt to check the accuracy of the other methods, as has been done with satisfactory results for the upper ionosphere in conjunction with ordinary ionosondes.

Two nonrocket radio probing methods are reasonably unambiguous. One, due to Fejer (1955), measures wave interaction in the ionosphere, but requires large radiated power. The other, by Gardner and Pawsey (1955) has the very interesting feature that, instead of the results being uncertain because of the presence of fine structure, the method actually requires and uses partial reflections from the region to deduce the electron density at various heights. To lend emphasis to this method, which has not been fully exploited in lower ionosphere studies, a brief description follows.

Gardner and Pawsey found that the polarization of echoes returned from the *D*-region was elliptical, but with the sense of rotation of an extrordinary wave. This was compatible with partial reflection from "under-dense" clouds of ionization for a frequency above the gyrofrequency. Upon examining the magnetoionic formula for the refractive index they found that, to a good approximation, the ratio of ordinary to extraordinary polarization of the partial reflection coefficient was almost solely a function of the collision frequency at the reflecting region. The differential absorption between the two components of the wave from a reflecting region was almost solely a function of the electron density at each element of the path below the region for a fixed collision frequency. Therefore, by assuming a reasonable variation of collision frequency with height, the effect of the partial reflection at the discontinuity could be removed from the resultant received wave, and a curve of differential absorption against height could be plotted by measuring the polarization of the received wave from all heights which returned usable echoes. The slope of that curve, together with the value of collision frequency at each point, provided a measure of electron density at the various heights.

An assumption of the values of collision frequency was necessary, as mentioned above, but Gardner and Pawsey were also able to measure it near the bottom of the reflecting region by assuming no absorption had affected the shortest range echoes. The polarization measurement then represented only the ratio of the partial reflection coefficients which, as noted above, resulted in a determination of the collision frequency at that level. An extrapolation of this collision frequency to greater heights was done by assuming an arbitrary law of its variation.

Specific Recommendations for Future Radio Experiments

1. Confirmation of the accuracy of the Gardner and Pawsey polarization method is desirable before it is given extended use. This confirmation could be achieved by a Gardner and Pawsey experiment combined with simultaneous rocket electron profile measurements. It is very necessary to determine whether or not all concentration "blobs" of electrons are of less than critical density, particularly as Gregory at times obtains a high reflection coefficient which might indicate an over-dense condition. Such a condition would affect the polarization and invalidate the method. An alternate method of checking would be to duplicate the Gardner and Pawsey method at two reasonably close frequencies both a little above the gyrofrequency. If the same profile is obtained at the two frequencies, then the electron density of the scattering blobs must be considerably below the critical density.

2. A single MF vertical pulse equipment should be run continuously—without any inserted attenuation—through the dawn, daytime and dusk period to look for any systematic height change associated with a specific stratum. This should give evidence bearing on the photoemissive origin or otherwise of the reflective region.

3. The synoptic behavior of all the lower strata is virtually an unexplored field. Since VHF forward scatter communication has been shown to be directly related to these strata, seasonal and geographic studies are necessary to allow the optimum design of forward scatter aerials.

4. The resolution of the height and width of strata by vertical pulse techniques is dependent on the aerial beam width. It would be desirable either to alter a beam width, or to use two experiments together—say 2 and 40 Mc—with considerably different beam widths, to determine the extent of range spread introduced by off-vertical directions. Once this had been determined, experiments could then be designed to examine particular heights for horizontal motion at low heights, and to investigate the relationship between turbulence and height.

5. Further studies are required of the apparent lowering of strata heights during magnetically disturbed periods. Again rocket flights at such a period, together with quiet period control flights, should reinforce the vertical pulse techniques in determining the electron profiles at such times.

6. It would appear that the 95-km lower *E*-ledge is created by the action of Lyman β radiation on O_2 (Houston 1958), and the 83-km region is attributed to Lyman α. If it is assumed that the intensities of both solar radiations change together, then some parallelism should be evident in the behavior of the two regions. Parallelism of behavior of *any* lower region with the 95-km daytime region would be evidence favoring a photoemissive origin of the lower region.

14. References

Appleton, E. V., On some measurements of the equivalent height of the atmospheric ionized layer, Proc. Roy. Soc. A. **126**, 542 (1930).

Appleton, E. V., Naismith, K., and Ingram, L. J., The critical frequency method of measuring upper atmospheric ionization, Proc. Phys. Soc. (London) **51**, 81 (1939).

Appleton, E. V. and Piddington, J. H., The reflection coefficients of ionospheric regions, Proc. Roy. Soc. A. **164**, 467 (1938).

Appleton, E. V. and Piggott, W. R., Ionospheric absorption measurements during a sunspot cycle, J. Atmospheric Terrest. Phys. **5,** 141 (1954).

Bailey, D. K., Disturbances in the lower ionosphere observed at VHF following the solar flare of 23 February 1956 with particular reference to auroral zone absorption, J. Geophys. Research **62,** 431 (1957).
Bailey, D. K., Bateman, R. and Kirby, R. C., Radio propagation at VHF by scattering and other processes in the lower ionosphere, Proc. IRE **43,** 1,181 (1955).
Bain, W. C., Observations on the propagation of very long radio waves reflected obliquely from the ionosphere during a solar flare, J. Atmospheric Terrest. Phys. **3,** 141 (1953).
Bain, W. C. and Bracewell, R. N., An explanation of radio propagation at 16 kc in terms of two layers below *E* Layer, J. Atmospheric Terrest. Phys. **2,** 216 (1952).
Banerji, R. B., Heights of irregularities giving rise to the fading of 150 kc waves, J. Geophys. Research **60,** 431 (1955).
Bates, D. R., The formation of the *E*-layer, Proc. Mixed Commission on the Ionosphere, **82,** (Brussels 1954).
Bates, D. R., Recombination in the ionosphere, Solar Eclipses and the Ionosphere, Spec. supp. to J. Atmospheric Terrest. Phys. **6,** 191 (1956).
Bates, D. R., Formation of the ionized layers, Solar Eclipses and the Ionosphere, Spec. supp. to J. Atmospheric Terrest. Phys. **6,** 184 (1956).
Bates, D. R. and Dalgarno, A., The altitudes of the luminous layers of the earth's atmosphere, J. Atmospheric Terrest. Phys. **4,** 122 (1953).
Bates, D. R. and Massey, H. S. W., The negative ion concentration in the lower ionosphere, J. Atmospheric Terrest. Phys. **2,** 1 (1951).
Benner, A. H., Grace, C. H., and Kelso, J. M., Polarization of low frequency radio waves reflected from the ionosphere, Proc. IRE **38,** 951 (1950).
Berg, O. E., Koomen, M., Meredith, L., and Scolnik, R., The altitude of the (OI) 5577 A line in the night airglow measured from a rocket, J. Geophys. Research **61,** 302 (1956).
Bigg, E. K., Detection of atmospheric dust and temperature inversions by twilight scattering, Nature **177,** 77 (1956).
Booker, H. G., Turbulence in the ionosphere with application to meteor trails, radio star scintillations, auroral radio echoes, and other phenomena, J. Geophys. Research **61,** 673 (1956).
Bowhill, S. A., The fading of radio waves of frequencies between 16 and 2,400 kc, J. Atmospheric Terrest. Phys. **8,** 129 (1956).
Bowhill, S. A., Ionospheric irregularities causing random fading of very low frequencies, J. Atmospheric Terrest. Phys. **11,** 91 (1957).
Bowles, K. L., Observation of vertical incidence scatter from the ionosphere at 41 mc, Phys. Rev. Letters **1,** 454 (1958).
Bracewell, R. N., Theory of formation of an ionospheric layer below *E* layer based on eclipse and solar flare effects at 16 kc/sec, J. Atmospheric Terrest. Phys.**2,** 226 (1952).
Bracewell, R. N., Budden, K. G., Ratcliffe, J. A., Straker, T. W., and Weekes, M. A., The ionospheric propagation of low and very low frequency radio waves over distances less than 1,000 km, Proc. IEE Part III **98,** 221 (1951).
Briggs, B. H., An investigation of certain properties of the ionosphere by means of a rapid frequency change experiment, Proc. Phys. Soc. B. **64,** 255 (1951).
Briggs, B. H. and Spencer, M., Horizontal movements in the ionosphere, Rep. on Prog. in Phys. **17,** 245 (1954).
Budden, K. G., The propagation of VLF radio waves to great distances, Phil. Mag. **44,** 504 (1953).
Budden, K. G., The numerical solution of differential equations governing reflection of long radio waves from the ionosphere, Proc. Roy. Soc. A. **227,** 516 (1955).

Caton, P. G. F. and Pierce, E. T., The waveforms of sferics, Phil. Mag. **43,** 393 (1952).
Chamberlain, J. W., Hunten, D. M. and Mack, J. E., Resonance scattering by atmospheric Na.—IV, J. Atmospheric Terrest. Phys. **12,** 153 (1958).
Chapman, S., The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth, Proc. Phys. Soc. **43,** 26 (1931).
Chubb, T. A., Friedman, H., Kreplin, R. W., and Kupperian, J. E. Jr., Rocket observation of X-ray emission in a solar flare, Nature **179,** 861 (1957).

Davids, N. and Parkinson, R. W., Wave solutions for critical and near critical coupling conditions in the ionosphere, J. Atmospheric Terrest. Phys. **7,** 173 (1955).
Deb, S., Origin of the *D* layer, Indian J. Phys. **14,** 89 (1940).
Dieminger, W., On the causes of excessive absorption in the ionosphere on winter days, J. Atmospheric Terrest. Phys. **2,** 340 (1952).
(See also Dieminger, W. and Hofmann-Heyden, A. E. Naturwissenshaften **39,** 84 (1952).
Dubin, M., Meteor ionization in the *E* region, Meteors, Spec. supp. to J. Atmospheric Terrest. Phys. **2,** 111 (1955).

Eckersley, T. L., Observations of scatter clouds, Nature **162,** 24 (1948).
Edwards, H. D., Bedinger, J. F., Manring, E. R., and Cooper, C. D., Emission from a sodium cloud artificially produced by means of a rocket, The Airglow and the Aurorae (Pergamon Press, 122, 1955).
Ellyett, C. D., Echoes at *D* heights with special reference to the Pacific Islands, Terrestrial Magnetism and Atmospheric Elec. **52,** 1 (1947).
Elsasser, H. and Siedentopf, H., The height of emission of the 5577 line of the airglow as observed on the Jungfraujoch, J. Atmospheric Terrest. Phys. **8,** 222 (1956).
Evans, S., Scale heights and pressures in the upper atmosphere from radio echo observations of meteors, Monthly Notices Roy. Astron. Soc. **114,** 63 (1954).
Evans, S., Atmospheric pressures and scale heights from radio echo observations of meteors, Meteors, Spec. supp. to J. Atmospheric Terrest. Phys. **2,** 86 (1955).

Fejar, J. A., The interaction of pulsed radio waves in the ionosphere, J. Atmospheric Terrest. Phys. **7,** 322 (1955).
Fleagle, R. G., Inferences concerning the dynamics of the mesopause, J. Geophys. Research **63,** 137 (1958).
Forsyth, P. A. Currie, B. W., and Vawter, F. E., Scattering of 56 mc radio waves from the lower ionosphere, Nature **171,** 35 (1953).
Friedman, H. and Chubb, T. A., Solar X-ray emission and the height of the *D* layer during radio fadeout, The Physics of the Ionosphere (Pergamon Press, 58, 1955).

Gallet, R. M., The nature of the sporadic *E* layer and turbulence in the upper atmosphere, C. R. Acad. Sci. **233,** 1649 (1951).
Gallet, R. M., Aerodynamical mechanisms producing electronic density fluctuations in turbulent ionized layers, Proc. IRE **43,** 1240 (1955).
Gardner, F. F. and Pawsey, J. L., Study of the ionospheric *D* region using partial reflections, J. Atmospheric Terrest. Phys. **3,** 321 (1953).
Gibbons, J. J. and Nertney, R. J., Wave solutions, including coupling, of ionospherically reflected long radio waves for a particular *E* region model, J. Geophys. Research **57,** 323 (1952).
Gnanalingam, S. and Weekes, K., Weak echoes from the ionosphere with radio waves of frequency 1.42 mc, Nature **170,** 113 (1952).
Gnanalingam, S. and Weekes, K., *D* region echoes observed with a radio wave of frequency 1.4 mc, The Physics of the Ionosphere (Pergamon Press, 1955).
Greenhow, J. S., Characteristics of radio echoes from meteor trails: III. The behaviour of the electron trails after formation, Proc. Phys. Soc. B, **65,** 169 (1952).
Greenhow, J. S., Systematic wind measurements at altitudes of 80–100 km using radio echoes from meteor trails, Phil. Mag. **45,** 71 (1954).
Greenhow, J. S. and Neufeld, E. L., Diurnal and seasonal wind variations in the upper atmosphere, Phil. Mag. **46,** 549 (1955).
Greenhow, J. S. and Neufeld, E. L., The height variation of upper atmospheric winds, Phil. Mag. (8th Ser.) **1,** 1157 (1956).
Gregory, J. B., Ionospheric reflections from heights below the *E* region, Australian J. Phys. **9,** 324 (1956).

Gregory, J. B., The relation of forward scattering of very high frequency radio waves to partial reflection of medium frequency waves at vertical incidence, J. Geophys. Research **62,** 383 (1957).
Gregory, J. B. (a), The mesopause region of the ionosphere, Nature **181,** 753 (1958).
Gregory, J. B. (b), Medium frequency observations of the lower ionosphere during sudden disturbances, J. Geophys. Research **63,** 273 (1958).
Gregory, J. B., (Private communication 1959).

Helliwell, R. A., Ionospheric virtual height measurements at 100 kc, Proc. IRE **37,** 887 (1949).
Helliwell, R. A., Sporadic *E* stratification and correlation with low frequency soundings, Trans. IRE, PGAP **3,** 140 (1952).
Helliwell, R. A., Mallinkrodt, A. J., and Kruse, F. W. Jr., Fine structure of the lower ionosphere, J. Geophys. Research **56,** 53 (1951).
Hepburn, F. and Pierce, E. T., Atmospherics with VLF components, Nature **171,** 837 (1953).
Heppner, J. P. and Meredith, L. H., Nightglow emission altitudes from rocket measurements, J. Geophys. Research **63,** 51 (1958).
Herlofson, N., The theory of meteor ionization, Rep. on Prog. in Phys. **11,** 444, (1948).
Herlofson, N., Plasma resonance in ionospheric irregularities, Arkiv För Fysik **3,** 247 (1951). (See also Tonks, L., Phys. Rev. **38,** 1219, 1931)
Hey, J. A. and Stewart, G. S., Radar observations of meteors, Proc. Phys. Soc. **59,** 858 (1947).
Hoffmeister, C., Cosmic influence on the upper strata of the atmosphere, Naturwiss. **22,** 458 (1934).
Hopkins, H. G. and Reynolds, L. G., An experimental investigation of short distance ionospheric propagation at low and very low frequencies, Proc. IEE Part III **101,** 21 (1954).
Houston, R. E., The effect of certain solar radiations on the lower ionosphere, J. Atmospheric Terrest. Phys. **12,** 225 (1958).
Hunten, D. M., Season variations of twilight Na emission, The Airglow and the Aurorae, Spec. Supp. to J. Atmospheric Terrest. Phys. **5,** 114 (1955).
Hunten, D. M. and Shepherd, G. G., A study of sodium at twilight: II Observations on the distribution, J. Atmospheric Terrest. Phys. **5,** 57 (1954).

Jackson, J. E., Measurements in the *E* layer with the Navy Viking rocket, J. Geophys. Research **59,** 377 (1954).
Jackson, J. E. and Seddon, J. C., Ionosphere electron density measurements with the Navy Aerobee-Hi rocket, J. Geophys. Research **63,** 197 (1958).
Jones, A. V., CaII emission lines in the twilight spectrum, Nature, **178,** 276 (1956).
Jones, A. V., Calcium and oxygen in the twilight airglow, Annal. die Geophysique **14,** 179 (1958).
Jones, I. L., The height variation of drift in the *E* Region, J. Atmospheric Terrest. Phys. **12,** 68 (1958). (See also special Supplement to J. Atmospheric Terrest. Phys. **12,** Part II, 20, 1957).
Jones, R. E., Millman, G. H., and Nertney, R. J., The heights of ionospheric winds as measured at long radio wavelengths, J. Atmospheric Terrest. Phys. **3,** 79 (1953).

Kaiser, T. R., Radio echo studies of meteor ionization, Phil. Mag. Supp. **2,** 495 (1953).
Kallman, H. K., Electron distribution in a new model of the ionosphere, Special Supp. J. Atmospheric Terrest. Phys. **82** (1957) *Pergamon Press.*
Khvostikov, I. A. and Sevcenko, A. N., Application of the polarimetric method in the study of the structure of the upper layers of the atmosphere, Compt. rend. acad. sci. U.R.S.S. **IV,** No. 8, 359 (1936).
Kilpatrick, E. L. (Private communication 1957).
Kitchen, F. A., Pressey, B. G., and Tremellen, K. W., A review of present knowledge of the ionospheric propagation of very-low, low, and medium frequency waves, Proc. IEE, Pt. 3 **100,** 100 (1953).

Landmark, B., Echoes from the lower ionosphere during polar blackouts, J. Atmospheric Terrest. Phys. **12,** 79 (1958).
Lien, J. R., Marcou, R. J., Ulwick, J. C., McMorrow, D. R., Linsford, D. B., and Haycock, O. C., Bifurcation in the *E* region, Phys. Rev. **92,** 508 (1953).
Liller, W. and Whipple, F. L., High altitude winds by meteor-train photography, Rocket Exploration of the Upper Atmosphere. Spec. Supp. to J. Atmospheric Terrest. Phys. **I,** 112 (1954).
Lindquist, R., An interpretation of vertical incidence equivalent height versus time recordings at 150 kc, J. Atmospheric Terrest. Phys. **4,** 10 (1953).
Ludlam, F. H., Noctilucent clouds, Tellus **9,** 341 (1957).

Manning, L. A., Peterson, A. M., and Villard, O. G., Ionospheric wind analysis by meteoric echo techniques, J. Geophys. Research **59,** 47 (1954).
Manring, E. R. and Pettit, H. B., Observations of the OI 5577 nightglow at Sacramento Peak, New Mexico, February through June 1955, The Airglow and the Aurorae, Supp. to J. Atmospheric Terrest. Phys. **5,** 29 (1955).
Maxwell, A., Turbulence in the upper ionosphere, Phil. Mag. **45,** 1,247 (1954).
McKinley, D. W. R. and Millman, P. M., A phenomenological theory of radar echoes from meteors, Proc. IRE **37,** 364 (1949).
McKinley, D. W. R. and Millman, P. M., Long duration echoes from the aurora, meteors and the ionosphere, Can. J. Phys. **31,** 171 (1953).
Millman, P. M. (Private communication 1957).
Millman, P. M and McKinley, D. W. R., Three-station radar and visual triangulation of meteors, Sky and Telescope, 8 (March 1949).
Mitra, A. P., The *D*-layer of the ionosphere, J. Geophys. Research **56,** 373 (1951).
Mitra, A. P., A tentative model of the equilibrium height distribution of NO in the high atmosphere and the resulting *D*-layer, J. Atmospheric Terrest. Phys. **5,** 28 (1954).
Mitra, A. P. and Jones, R. E., Recombination in the lower ionosphere, J. Geophys. Research **59,** 391 (1954).
Murgatroyd, R. J., Winds and temperatures between 20 km and 100 km—a review, J. Roy. Met. Soc. **83,** 417 (1957).

Naismith, R. and Bramley, E. N., Time delay measurements on radio transmissions, Wireless Engr. **28,** 271 (1951).
Nertney, R. J., Comments concerning the paper by Helliwell et al., J. Geophys. Research **56,** 449 (1951).
Nertney, R. J., The lower *E* and *D*-region of the ionosphere as deduced from long radio wave measurements, J. Atmospheric Terrest. Phys. **3,** 92 (1953).
Nicolet, M., Action of solar radiation in the upper atmosphere, Annal. die Geophysique **8,** 141 No. 2 (1952).
Nicolet, M., The *D* layer and nitric oxide, Proc. Mixed Commission on the Ionosphere, Brussels —, 40 (1954).

Öpik, E. J., Problems in the physics of meteors, Amer. J. Phys. **26,** 70 (1958).

Parkinson, R. W., The night-time lower ionosphere as deduced from a theoretical and experimental investigation of coupling phenomena at 150 kc/s, J. Atmospheric Terrest. Phys. **7,** 203 (1955).
Parkinson, R. W., The fading periods of the *E*-region coupling echo at 150 kc/s, J. Atmospheric Terrest. Phys. **8,** 158 (1956).
Paton, J., Auroras and luminous night clouds, Nature **168,** 487 (1951).
Pfister, W. and Ulwich, J. C., The analysis of rocket experiments in terms of electron-density distributions, J. Geophys. Research **63,** 315 (1958).
Piddington, J. H., The modes of formation of the ionospheric layers, J. Geophys. Research **56,** 409 (1951).
Pineo, V. C., Oblique incidence measurements of the heights at which ionospheric scattering of VHF radio waves occurs, J. Geophys. Research **61,** 165 (1956).
Poeverlein, H., Low frequency reflection in the ionosphere—I, J. Atmospheric Terrest. Mag. **12,** 126 (1958); Low frequency reflection in the ionosphere—II, J. Atmospheric Terrest. Mag. **12,** 236 (1958).
Porter, J. G., An analysis of British meteor data, Mon. Not. Roy. Astron. Soc. **104,** 257 (1944).

Prentice, J. P. M., Meteors, Rep. on Prog. in Phys. **XI,** 389 (1948).

Ratcliffe, J. A. and White, E. L. C., The effect of the earth's magnetic field on the propagation of short wireless waves, Phil. Mag. **16,** 125 (1933).
Rawer, K., A critical survey of eclipse information on recombination data, Solar Eclipses and the Ionosphere, Spec. Supp. to J. Atmospheric Terrest. Mag. **6,** 198 (1956).
Rivault, R., The microstructure of atmospherics. Ionospheric and meteorological applications of those of type 4, C. R. Acad. Sci. **230,** 1846 (1950).
Roach, F. E., The night airglow, Proc. IRE **47,** 267 (1959).
Roach, F. E. and Meinel, A. B., The height of the night glow by the van Rhyn method, Astrophys. J. **122,** 530 (1955).
Roach, F. E., Megill, L. R., Rees, M. H., and Marovich, E., The height of night glow 5577; J. Atmospheric Terrest. Phys. **12,** 171 (1958).
Roach, F. E., Tandberg-Hansen, E., and Megill, L. R. (a), The characteristic size of airglow cells, J. Atmospheric Terrest. Phys. **13,** 113 (1958).
Roach, F. E., Tandberg-Hansen, E., and Megill, L. R. (b), Movements of airglow cells, J. Atmospheric Terrest. Phys. **13,** 122 (1958).

Seddon, J. C., Differential absorption in the *D* and lower *E* regions, J. Geophys. Research **63,** 209 (1958).
Seddon, J. C., Pickar, A. D., and Jackson, J. E., Continuous electron density measurements up to 200 km, J. Geophys. Research **59,** 513 (1954).
St. Amand, P., Pettit, H. B., Roach, F. E., and Williams, D. R., On a new method of determining the height of the airglow, J. Atmos. Terr. Phys. **6,** 189 (1955).
Straker, T. W., The ionospheric propagation of radio waves of frequency 16 kc/s over short distances, Monograph No. 114, Inst. Elec. Engrs. (London, 1955).

Thomas, J. A. and Burke, M. J., Motion of the night-time Es region at Brisbane, Australian J. Phys. **9,** 440 (1956).

Wait, J. R., The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (1957).
Wait, J. R., An extension to the mode theory of VLF Ionospheric propagation, J. Geophys. Research **63,** 125 (1958).
Wait, J. R. and Perry, L. B., Calculations of ionospheric reflection coefficients at very low frequencies, J. Geophys. Research **62,** 43 (1957).
Watanabe, K., Marmo, F., and Inn, E. C. Y., Formation of the *D* layer, Phys. Rev. **90,** 155 (1953).
Watanabe, K., Marmo, F. F., and Pressman, J., Formation of the lower ionosphere, J. Geophys. Research **60,** 513 (1955).
Watts, J. M. (a), A note on the polarization of low frequency ionosphere echoes, J. Geophys. Research **57,** 287 (1952).
Watts, J. M. (b), Oblique incidence propagation at 300 kc using the pulse technique, J. Geophys. Research **57,** 487 (1952).
Watts, J. M. and Brown, J. N., Effects of ionosphere disturbances on low frequency propagation, J. Geophys. Research **56,** 403 (1951).
Watts, J. M. and Brown, J. N., Some results of sweep frequency investigation in the low frequency band, J. Geophys. Research **59,** 71 (1954).
Waynick, A. H., The lowest ionosphere, Physics of the Ionosphere, Physical Soc. (1955).
Weekes, K., The ground interference pattern of very low frequency radio waves, Proc. IEE, Part III **97,** 100 (1950).
Weekes, K., The physical state of the upper atmosphere, J. Roy. Meteorol Soc. **80,** 2 (1954).
Weiss, A. A., Radio echo observations of meteors in the southern hemisphere, Meteors, Spec. Supp. to J. Atmospheric Terrest. Phys. II 195 (1955).
Whale, H. A., Fine structure of the ionospheric region *E*, J. Atmospheric Terrest. Phys. **1,** 233 (1951).
Wheelon, A. D., Diurnal variations of signal level and scattering heights for VHF propagation, J. Geophys. Research **62,** 255 (1957).
Whipple, F. L., Results of rocket and meteor research, Bull. Am. Meteorol. Soc. **33,** 13 (1952).

BOULDER, COLO. (Paper 63D2-13)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Effect of Small Irregularities on the Constitutive Relations for the Ionosphere

K. G. Budden[1]

(March 18, 1959)

Irregularities in the ionosphere which are small compared with one wavelength may modify the constitutive relations, and hence, may affect the refractive indices for electromagnetic waves. The modifications are in some ways similar to those which would be introduced into the Appleton-Hartree formula by a Lorentz force. The theory is given first for the case when the irregularities extend only in one dimension, and it is found that even in a loss-free medium the refractive index now has an imaginary part which might be associated with loss of energy from the wave by scattering. The theory for three-dimensional irregularities is then discussed but is more difficult, and a method of successive approximations is used. The results indicate that small irregularities may play an important part in the propagation of very-low-frequency radio waves in the ionosphere. In particular, they may explain why "whistlers" are observed only on comparatively rare occasions.

1. Introduction

In one method of deriving the formulas of the magneto-ionic theory [1],[2] the ionized medium is assumed to have an electric polarization **P** which is related to the electric field **E** through a set of equations called the "constitutive relations." By using Maxwell's equations together with the constitutive relations, the Appleton-Hartree formula for the refractive index may be deduced. The vector **P** is given by:

$$\mathbf{P} = N e \mathbf{r} \tag{1}$$

where N is the number of electrons per unit volume, e is the charge on one electron, and **r** is the average displacement of an electron from its mean position. **P** is thus defined by taking an average over a volume large enough to contain many electrons. When used in Maxwell's equations it is assumed to represent a vector field, continuously distributed in space, and approximately constant over distances which are small compared to a wavelenth. There would be no meaning in speaking of the value of **P** at a specific point in the free space between the electrons. The use of Maxwell's equations implies a "smoothing out" process over a distance which must be large compared with $N^{-1/3}$.

There is now much evidence to show that the ionosphere is an irregular medium, so that the electron-number density N varies from place to place. The effect of these irregularities on the propagation of waves through the medium may be treated by considering the energy scattered from them, or by studying their effect on the refraction and diffraction of the waves. In the present paper, an alternative method is suggested for studying this problem in the special case when the irregularities are very small compared with one wavelength. The electric field **E** and the polarization **P** depend on the electron-number density N and vary from place to place because N varies, but it is possible to find their average values, $\mathbf{E}_0$ and $\mathbf{P}_0$, by a "smoothing out" process similar to that already mentioned. It is the average values which must be used in Maxwell's equations; and, in order to find the refractive index, it is necessary to know how $\mathbf{P}_0$ depends on $\mathbf{E}_0$. A new set of constitutive relations is therefore required giving $\mathbf{P}_0$ in terms of $\mathbf{E}_0$,

[1] Cavendish Laboratory, Cambridge, England. This work was done at the National Bureau of Standards while the author was on sabbatical leave from Cambridge University. The work was supported by the U.S. Air Force and the Bureau.

[2] Figures in brackets indicate the literature references at the end of this paper.

and this is, in general, different from the more familiar relations which give **P** in terms of **E**. Hence, the refractive index must also be modified when small irregularities are present. The purpose of the present paper is to find these modifications in some special cases.

2. Qualitative Description of the Method

The ionosphere is an electrically neutral medium, and it is assumed that the number of heavy positive ions per unit volume is the same as the number of electrons, N. Variations in N from place to place, therefore, occur equally for electrons and positive ions. They are assumed to be statistically stationary, and the average value of N over a very large region is N_0.

Suppose that a uniform electric field, **E**, is applied to the ionosphere. If **E** varies harmonically in time, the electron displacements also vary harmonically with the same frequency, and at any one instant all the electrons are displaced the same distance, **r**, from their mean positions The positive ions are assumed to be so heavy that they do not move. Let x be distance measured in the direction of **E**, and suppose that N varies with x as shown in figure 1. When the electrons are displaced, the effect is to introduce some space charge, which is negative where N is decreasing in the direction of **r**, and positive where it is increasing. The first effect of the space charge is to introduce an additional electric field, $\Delta\mathbf{E}$, whose average value is zero as long as the electrons are all displaced by the same amount. To a first approximation, therefore, the average value, $\mathbf{E}_0$, of the electric field is just the original field **E**, and the average value of the polarization is $\mathbf{P}_0 = N_0 e\mathbf{r}$. In this case, $\mathbf{P}_0$ and $\mathbf{E}_0$ are related by the same constitutive relations as for a homogeneous ionosphere. This results from the assumption that the electron displacements **r** are all the same.

The additional electric field $\Delta\mathbf{E}$ exerts a force on the electrons which changes their displacements. This in turn changes the space charge distribution, which then affects $\Delta\mathbf{E}$. This is a second-order effect since $\Delta\mathbf{E}$ is small, but it will be shown that it causes both **E** and **P** to have new average values, $\mathbf{E}_0$ and $\mathbf{P}_0$, which are related in a different way from the relation between **E** and **P**.

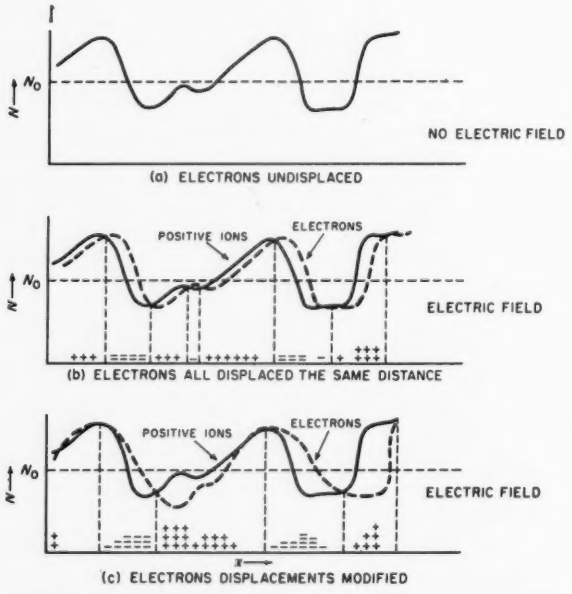


FIGURE 1. *When electrons are displaced in an irregular ionized medium, space charge appears and gives rise to electric forces which modify the displacements.*

3. Comparison With the Lorentz Force

When the magneto-ionic theory was first formulated, it was believed that the electric force on an electron in the ionosphere was made up of two parts. One part, $e\mathbf{E}$, arose from the electric field $\mathbf{E}$ impressed on the medium and the other part, $1/3\ \mathbf{P}e$ (rationalized units), arose from the polarization of the medium itself, and is called the Lorentz force. For many crystalline dielectrics, a Lorentz force must certainly be included, but there is now evidence that in the ionosphere it should be omitted. The strongest evidence is from the theory of "whistlers" which consist of very-low-frequency waves propagated in the extensive ionized region surrounding the earth, in directions close to the lines of force of the earth's magnetic field [5]. The Lorentz force would so modify the constitutive relations for the ionosphere, that these low-frequency waves could not be propagated.

The modification of the constitutive relations that would be introduced by small irregularities is similar in some ways to that which would arise from the Lorentz force. Both effects occur because there is a force acting on the electrons which arises from the polarization $\mathbf{P}$. For the Lorentz force this is simply proportional to $\mathbf{P}$, but for small irregularities it depends on both $\mathbf{P}$ and $\mathbf{E}$. For this reason, the effect of small irregularities is more complicated than that of a Lorentz force, and it will be shown that the additional force on the electron may depend upon the state of polarization of the wave. For circularly polarized waves of very low frequency traveling in the direction of the earth's magnetic field (whistlers), the effect of irregularities may have the same direction as the Lorentz force. This affects the dispersion law for these waves, and if the irregularities are large enough in amplitude, there is a lower limit to the frequencies that can be propagated.

4. Definitions of E and D

For a medium which is not free space, the electric intensity $\mathbf{E}$ is usually defined in terms of a long thin cavity. The component of $\mathbf{E}$ in a given direction is found as follows. A long thin cavity is imagined to be cut in the medium parallel to the given direction. Its length must be so small that the electric state of the medium does not change appreciably within it. For waves in a homogeneous ionosphere this means that the length is small compared to a wavelength. An infinitesimal test charge δq is placed at the center of the cavity and the component $\delta\mathbf{F}$ of the force acting on it in the direction of the cavity is measured. The component of $\mathbf{E}$ in this direction is

$$\lim_{\delta q \to 0} \delta\mathbf{F}/\delta q. \tag{1}$$

For a medium which contains discrete ions and electrons, this definition can still be used provided that the cross section of the cavity is very large compared with $N^{-\frac{1}{2}}$. It is the electric intensity so defined which is to be used in Maxwell's equations. The cavity definition thus effects the "smoothing out" process mentioned in the introduction. A similar argument applies to the electric displacement $\mathbf{D}$ which is defined in terms of the force normal to the plane of a flat, plate-like cavity on a test charge in the cavity. The longest dimension of the cavity must be small compared to a wavelength, λ and the thickness must be large compared with $N^{-\frac{1}{3}}$.

Now consider an ionized medium in which there are irregularities small in size compared to one wavelength. Suppose that the size of the irregularities is given by a length, l, which may be thought of as the average distance between one point of maximum electron density and its nearest neighbor. Then $l << \lambda$. The electric intensity can now be defined in two ways.

The "local" electric intensity $\mathbf{E}$ is defined in terms of a long thin cavity whose length is very small compared with l, and whose cross section is large compared with $N^{-\frac{1}{2}}$. Its value depends on where the cavity is placed in relation to the maximums of N, and is affected by the local configuration of the space charge. It is this value of $\mathbf{E}$ which determines the force on an electron.

The average electric intensity $\mathbf{E}_0$ is defined in terms of a long thin cavity whose length is very small compared with λ, and whose cross section is large compared with l. In this case the distance of the test charge from the walls of the cavity is very large compared with l, so that the value of $\mathbf{E}_0$ is unaffected by local configuration of the irregularities. It is $\mathbf{E}_0$ which must be used in Maxwell's equations. The cavity definition here effects the "smoothing out" process over a distance small compared to a wavelength but large compared with l.

In a similar way, flat plate-like cavities may be used to define the local value $\mathbf{D}$ and the average value $\mathbf{D}_0$ of the electric displacement. The local electric polarization $\mathbf{P}$ is given by (1). Its average value is denoted by $\mathbf{P}_0$, and is *not*, in general, equal to $N_0 e \mathbf{r}_0$ where $\mathbf{r}_0$ is the average value of $\mathbf{r}$. The following relations may be derived by the standard methods of electromagnetic theory (rationalized units):

$$\mathbf{D}=\epsilon_0\mathbf{E}+\mathbf{P}, \tag{2}$$

$$\mathbf{D}_0=\epsilon_0\mathbf{E}_0+\mathbf{P}_0. \tag{3}$$

$\mathbf{P}$ and $\mathbf{E}$ are related by the same constitutive relations as for a homogeneous medium, but this is not true for $\mathbf{P}_0$ and $\mathbf{E}_0$.

Before attempting the general theory for three-dimensional space, the theory will be illustrated by discussing a simpler problem in one dimension.

5. Irregularities in One Dimension

Suppose that the electron density N is a function only of x, and consider the effect of an electric field parallel to the x direction. The effect of the earth's magnetic field is neglected. Our problem is to calculate the effective dielectric constant of the medium. The reader may find it helpful to imagine that a large sample of the medium is placed between two conducting plates in planes x=constant, forming a parallel plate condenser whose capacity is to be found. $\mathbf{D}$ and $\mathbf{E}$ are both parallel to the x direction and may be written simply as D and E. The electron number density N is given by

$$N=N_0+\Delta N, \tag{4}$$

where ΔN is small compared with N_0. Similarly the displacement r of an electron in the x direction from its mean position is given by

$$r=r_0+\Delta r. \tag{5}$$

Further

$$E=E_0+\Delta E. \tag{6}$$

Each of ΔN, Δr, and ΔE is a function of x, and has an average value of zero.

Now div $\mathbf{D}=0$ so that in this case $dD/dx=0$ and

$$D=D_0=\text{constant} \tag{7}$$

for all values of x. This equation automatically takes care of the effect of space charge. The space charge density is $\rho=-dP/dx$. It gives rise to an additional electric field which must be such that div $\mathbf{D}=0$. It would clearly be incorrect to write div $\mathbf{D}=\rho$, for this would include the space charge twice.

The electron displacement r is related to E thus.

$$r=\gamma E, \tag{8}$$

where γ is a constant which is, in general, complex. If the effect of collisions is neglected, $\gamma=-e/m\omega^2$ where m is the mass of an electron and ω is the angular wave frequency. If the

average number of collisions per unit time for an electron is ν, then $\gamma=e/(im\nu\omega-m\omega^2)$. Now $P=Ner=Ne\gamma E$ and hence

$$D=(\epsilon_0+Ne\gamma)E=D_0=\{\epsilon_0+(N_0+\Delta N)e\gamma\}(E_0+\Delta E). \tag{9}$$

This gives

$$\Delta E=\frac{D_0}{\epsilon_0+e\gamma N_0+e\gamma\Delta N}-E_0, \tag{10}$$

whose average value must be zero.

It is now necessary to make some further assumptions about the irregularities ΔN. Let $\Delta N=\xi$. It is assumed that ξ has a Gaussian distribution about zero, which means that the probability, $p(\xi)\,d\xi$, that ξ lies in the range ξ to $\xi+d\xi$ is given by

$$p(\xi)=\frac{1}{\sqrt{2}a\pi^{1/2}}\,e^{-\xi^2/2a^2}, \tag{11}$$

where

$$a^2=\overline{\Delta N^2}. \tag{12}$$

(A quantity with a bar over it denotes the average value of that quantity.) Then the average value of ΔE is

$$\overline{\Delta E}=\frac{1}{a\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{D_0}{\epsilon_0+e\gamma N_0+e\gamma\xi}\,e^{-\xi^2/2a^2}\,d\xi-E_0 \tag{13}$$

which must be zero. The integral is evaluated in appendix A. Let

$$\sqrt{2}ae\gamma=\zeta;\ \frac{\epsilon_0+e\gamma N_0}{\sqrt{2}ae\gamma}=\Omega. \tag{14}$$

Then the result of equating (13) to zero is:

$$\frac{E_0}{D_0}=-\frac{1}{\zeta}\left\{i\pi^{1/2}e^{-\Omega^2}-2F(\Omega)\right\}, \tag{15}$$

where

$$F(\Omega)=e^{-\Omega^2}\int_0^{\Omega}e^{u^2}du. \tag{16}$$

If electron collisions are neglected, γ is real and the integrand in (13) is then real at all points on the real ξ axis, but has a pole on this path. The principal value of the integral is then purely real. If γ has a small imaginary part, however, the integral has an imaginary part which does not tend to zero as $I(\gamma)\to 0$. Thus the limiting value of the integral is not the same as the principal value. Now it is impossible for the collision frequency to be exactly zero so that clearly the limiting value is the correct one to use when collisions are neglected.

Now

$$1+N_0e\gamma/\epsilon_0=\epsilon', \tag{17}$$

where ϵ' is the dielectric constant of a homogeneous medium with the electron number density N_0. The true dielectric constant is $\epsilon=D_0/\epsilon_0E_0$. Let

$$\overline{\left(\frac{\Delta N}{N_0}\right)^2}=\frac{a^2}{{N_0}^2}=\beta. \tag{18}$$

Then (14) gives

$$\zeta=(2\beta)^{1/2}\epsilon_0(\epsilon'-1);\ \Omega=(2\beta)^{-1/2}\epsilon'/(\epsilon'-1), \tag{19}$$

which may be substituted in (15).

If electron collisions are neglected:

$$\epsilon' = 1 - X, \text{ where } X = \frac{N_0 e^2}{\epsilon_0 m \omega^2}, \tag{20}$$

and (15) gives

$$\mu^2 = \epsilon = \sqrt{2} X \beta^{1/2} \{ i\pi^{1/2} e^{-\Omega^2} - 2F(\Omega) \}^{-1}. \tag{21}$$

This is the formula for the square of the refractive index in the one-dimensional case when small irregularities are present. It replaces the formula $\mu^2 = 1 - X$ which would be true for a homogeneous medium. The quantity X is the same as Appleton's [1] x and is proportional to the electron density. In figure 2 curves are plotted showing how the real and imaginary parts of μ depend upon X for various values of β.

Equation (21) shows that μ^2 has an imaginary part even when electron collisions are neglected. Hence, some energy would be lost from an electromagnetic wave although there is no physical mechanism for converting electric energy into heat. It is known that irregularities cause scattering of energy from the original wave, and this must account for the imaginary part of μ^2.

When $X \gg 1$, Ω is large since β is small. The term, $e^{-\Omega^2}$ may then be neglected, and $F(\Omega)$ may be replaced by the first two terms of its asymptotic expansion (eq A7, app. A). If powers of β higher than the first are neglected, the expression (21) becomes

$$\mu^2 \approx 1 - \frac{X}{1 + \beta X/(X-1)}, \tag{22}$$

which may be compared with the expression for μ^2 in a homogeneous medium when the Lorentz term is included, namely:

$$\mu^2 = 1 - \frac{X}{1 + (1/3)X}. \tag{23}$$

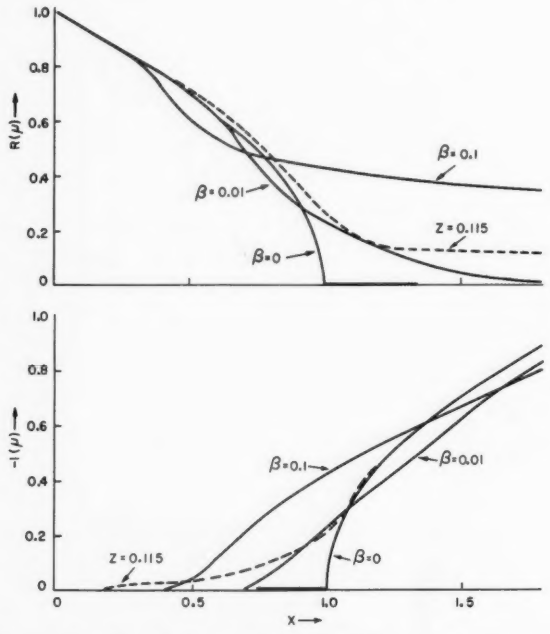


FIGURE 2. *The effect of small irregularities on the variation of $R(\mu)$ and $-I(\mu)$ with X.*

β is a measure of the strength of the irregularities. For the full curves electron collisions are neglected. For the dashed curves, $Z = \nu/\omega = 0.115$, and $\beta = 0$.

The factor $\beta/(X-1)$ in (22) is analogous to the Lorentz factor, 1/3, in (23). The effect of the irregularities has the same sign as a Lorentz force (since $X \gg 1$), but gets progressively smaller as X increases.

6. Irregularities in Three Dimensions, General Theory

There appears to be no direct way of extending the method of the last section to deal with irregularities in three dimensions. In one dimension the equation, div $\mathbf{D}=0$, has the simple integral $\mathbf{D}=$constant, but there is no correspondingly simple result for three dimensions. In this and the following sections, therefore, a solution by successive approximations will be attempted.

Cartesian coordinates x, y, z, are used. Let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ denote unit vectors parallel to the x, y, and z axes respectively. Then the position of a point (x, y, z) is given by the vector $\mathbf{s}=\mathbf{i}x+\mathbf{j}y+\mathbf{k}z$. The electron number density N is given by (4) where ΔN is now an irregular function of x, y, z. ΔN may be expressed as the sum of many Fourier terms thus:

$$\Delta N=\sum_K (\Delta N)_K, \tag{24}$$

$$(\Delta N)_K=A_K \cos{(\mathbf{K}\mathbf{s}-\phi_K)}, \tag{25}$$

where

$$\mathbf{K}=u\mathbf{i}+v\mathbf{j}+w\mathbf{k}. \tag{26}$$

The summation is really a triple sum over all values of u, v, w. A_K and ϕ_K are the amplitude and phase, respectively, of the Fourier term.

7. First Approximation

As a first approximation it is assumed that the electric field $\mathbf{E}$ in the medium is the same at all points and equal to $\mathbf{E}_1$. Then each electron is displaced from its mean position by the same amount $\mathbf{r}_1$ given by

$$\mathbf{r}_1=\mathbf{T}\cdot\mathbf{E}_1, \tag{27}$$

where $\mathbf{T}$ is a tensor which can be found from the equation of motion of an electron. Because of the irregularities ΔN, the displacement $\mathbf{r}_1$ results in the appearance of space charge which in turn gives an additional electric field $\Delta\mathbf{E}_1$. The contribution $\Delta\mathbf{E}_1{}^{(K)}$ to $\Delta\mathbf{E}_1$ from one Fourier component in (24) will now be found.

A single Fourier component $(\Delta N)_K$ is a one-dimensional variation of ΔN and may be treated as in section 5. The additional electric field $\Delta\mathbf{E}_1{}^{(K)}$ must be normal to the planes of constant ΔN for this component and so is parallel to the vector $\mathbf{K}$. At this stage $\mathbf{r}_1$ is taken as a constant, so that the average value of $\mathbf{E}_1$ must remain unaffected by the irregularities. Hence, $\overline{\Delta\mathbf{E}_1{}^{(K)}}=0$. For components parallel to $\mathbf{K}$ we have $\epsilon_0 E+P=D=$constant, so that $\epsilon_0 \Delta E+\Delta P=0$. Now $\Delta P=e(\Delta N)_K r_1{}^{(K)}$ where $r_1{}^{(K)}$ is the component of $\mathbf{r}_1$ parallel to $\mathbf{K}$. Hence

$$\Delta E_1{}^{(K)}=-\frac{e}{\epsilon_0} A_K \cos{(\mathbf{K}\mathbf{s}-\phi_K)} \frac{(\mathbf{r}_1\mathbf{K})}{\mathbf{K}^2}\cdot\mathbf{K}. \tag{28}$$

Here the factor $(\mathbf{r}_1\mathbf{K})/|K|$ is $r_1{}^{(K)}$, and the factor $\mathbf{K}/|K|$ is a unit vector in the direction of $\mathbf{K}$.

The total additional field $\Delta\mathbf{E}_1$ from the first approximation is given by:

$$\Delta\mathbf{E}_1=\sum_K \Delta\mathbf{E}_1{}^{(K)}. \tag{29}$$

8. Second Approximation

The effect of the appearance of the additional electric field $\Delta\mathbf{E}_1$ is to give the electrons an additional displacement,

$$\Delta\mathbf{r}_1=\mathbf{T}\cdot\Delta\mathbf{E}_1. \tag{30}$$

Both $\Delta\mathbf{r}_1$ and $\Delta\mathbf{E}_1$ are functions of position. $\Delta\mathbf{r}_1$ gives rise to further space charge which in turn results in a further contribution $\Delta\mathbf{E}_2$ to the electric field. The average value of $\Delta\mathbf{E}_1$ is zero, but the average value $\overline{\Delta\mathbf{E}_2}$ of $\Delta\mathbf{E}_2$ is not in general zero, and its value must now be found.

The resultant electron displacement is $\mathbf{r}_1+\Delta\mathbf{r}_1$, which must be substituted for $\mathbf{r}_1$ in eq (28). This gives:

$$\Delta\mathbf{E}_2{}^{(K)}=-\frac{e}{\epsilon_0}A_K\cos(\mathbf{K}\,\mathbf{s}-\phi_K)\frac{\{(\mathbf{r}_1+\Delta\mathbf{r}_1)\cdot\mathbf{K}\}}{\mathbf{K}^2}\cdot\mathbf{K}, \tag{31}$$

and $\Delta\mathbf{E}_2$ is given by

$$\Delta\mathbf{E}_2=\sum_K\Delta\mathbf{E}_2{}^{(K)}. \tag{32}$$

$\Delta\mathbf{r}_1$, and therefore $\Delta\mathbf{E}_2{}^{(K)}$, contain a summation over all values of $\mathbf{K}$, and so (32) contains a double summation. Now take the average of (32). Clearly $\cos(\mathbf{K}_1\mathbf{s}-\phi_{K_1})\cos(\mathbf{K}_2\mathbf{s}-\phi_{K_2})$ has the average value zero when $\mathbf{K}_1\neq\mathbf{K}_2$ and the average value 1/2 when $\mathbf{K}_1=\mathbf{K}_2$. The value of $\mathbf{r}_1$ may be inserted from (27). Hence:

$$\overline{\Delta\mathbf{E}_2}=\frac{e^2}{2\epsilon_0^2}\sum_K\frac{A_K^2\{\mathbf{T}\,\mathbf{K}\cdot\mathbf{K}\}\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}}{\mathbf{K}^4}\cdot\mathbf{K}. \tag{33}$$

The average electric field is now given by

$$\mathbf{E}_0=\mathbf{E}_1+\overline{\Delta\mathbf{E}_2}. \tag{34}$$

Next it is necessary to find the average value $\mathbf{P}_0$ of the polarization $\mathbf{P}$. From eq (1):

$$\mathbf{P}=e(N_0+\Delta N)(\mathbf{r}_1+\Delta\mathbf{r}_1), \tag{35}$$

and hence

$$\mathbf{P}_0=eN_0\mathbf{r}_1+e\,\overline{\Delta N\Delta\mathbf{r}_1}. \tag{36}$$

From (28), (29), and (30):

$$\Delta N\Delta\mathbf{r}_1=-\frac{e}{\epsilon_0}\sum_K A_K\cos(\mathbf{K}\mathbf{s}-\phi_K)\frac{(\mathbf{r}_1\cdot\mathbf{K})}{\mathbf{K}^2}\,\mathbf{T}\,\mathbf{K}\sum_K A_K\cos(\mathbf{K}\mathbf{s}-\phi_K), \tag{37}$$

and on taking the average:

$$\overline{\Delta N\Delta\mathbf{r}_1}=-\frac{e}{2\epsilon_0}\sum_K A_K^2\frac{(\mathbf{r}_1\cdot\mathbf{K})}{\mathbf{K}^2}\,\mathbf{T}\,\mathbf{K}. \tag{38}$$

Hence from (36) and (27)

$$\mathbf{P}_0=eN_0\mathbf{T}\,\mathbf{E}_1-\frac{e^2}{2\epsilon_0}\sum_K A_K^2\frac{\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}\mathbf{T}\,\mathbf{K}}{\mathbf{K}^2}. \tag{39}$$

In eq (33), (34), and (39), $\mathbf{P}_0$ and $\mathbf{E}_0$ are expressed in terms of $\mathbf{E}_1$. Elimination of $\mathbf{E}_1$ gives the constitutive relations between $\mathbf{P}_0$ and $\mathbf{E}_0$.

9. Discussion When T Is a Scalar

If the effect of the earth's magnetic field is neglected, the tensor $\mathbf{T}$ reduces to the scalar γ used in sec. 5. Then eq (33) becomes

$$\overline{\Delta\mathbf{E}_2}=\frac{\gamma^2e^2}{2\epsilon_0^2}\sum_K A_K^2\left\{\frac{(\mathbf{E}_1\cdot\mathbf{K})}{\mathbf{K}^2}\cdot\mathbf{K}\right\}. \tag{40}$$

The factor $[(\mathbf{E}_1\cdot\mathbf{K})/\mathbf{K}^2]\cdot\mathbf{K}$ is simply the component of $\mathbf{E}_1$ in the direction of $\mathbf{K}$.

It is now assumed that the irregularities are isotropic. This means that for a given value of $|K|$ the directions of the vectors, $\mathbf{K}$, are uniformly distributed in any hemisphere and the

probability of a given value of A_K is independent of the direction of K. Select first those values of $\mathbf{K}$ for which $|K|$ is in a specified range, $d|K|$, and A_K is in a specified range, dA_K, and perform the summation for these values first. Then the last factor in (36) is simply $1/2\ \mathbf{E}_1$. Hence (40) becomes:

$$\overline{\Delta \mathbf{E}_2}=\frac{1}{4}\frac{\gamma^2 e^2}{\epsilon_0^2}\mathbf{E}_1\sum_K A_K^2. \tag{41}$$

But

$$\frac{1}{2}\sum_K A_K^2=\overline{(\Delta N)^2}. \tag{42}$$

Hence (34), (41), and (42) give:

$$\mathbf{E}_0=\mathbf{E}_1\left\{1+\frac{1}{2}\frac{e^2\gamma^2 N_0^2}{\epsilon_0^2}\overline{\left(\frac{\Delta N}{N_0}\right)^2}\right\}. \tag{43}$$

Equation (39) becomes:

$$\mathbf{P}_0=\gamma e N_0\mathbf{E}_1-\frac{\gamma^2 e^2}{2\epsilon_0}\sum_K A_K^2\left\{\frac{(\mathbf{E}_1\cdot\mathbf{K})}{\mathbf{K}^2}\cdot\mathbf{K}\right\} \tag{44}$$

which reduces in a similar way to

$$\mathbf{P}_0=\mathbf{E}_1\left\{\gamma e N_0-\frac{1}{2}\frac{\gamma^2 e^2 N_0^2}{\epsilon_0}\overline{\left(\frac{\Delta N}{N_0}\right)^2}\right\}. \tag{45}$$

The dielectric constant ϵ is given by

$$\epsilon_0\,\epsilon\,\mathbf{E}_0=\epsilon_0\,\mathbf{E}_0+\mathbf{P}_0. \tag{46}$$

It is convenient to use the expression (17) for ϵ', and the abbreviation (18). Then:

$$\epsilon=\epsilon'/\{1+\tfrac{1}{2}\beta\,(\epsilon'-1)^2\}. \tag{47}$$

If electron collisions are neglected and (20) is used, eq (47) becomes:

$$\mu^2=\epsilon=\frac{1-X}{1+\frac{1}{2}\beta X^2} \tag{48}$$

which should be compared with the expressions (21) and (22) obtained for the one-dimensional case. The difference is not surprising since the plane sheets of space charge formed in the one-dimensional problem would be expected to have a larger and different effect from an irregular three-dimensional distribution of charge. It is probable that (48) is unreliable when X is near unity, for then the wave frequency is near the plasma frequency, and it is possible for the electrons to make oscillations of large amplitude. The approximations made in the derivation would then be invalid. This may also explain why the expression (48) is purely real. In a more accurate treatment it should have an imaginary part which is associated with loss of energy from the wave by scattering.

10. Discussion When T is a Tensor

The constitutive relations will now be derived for the special case of a plane wave with its wave normal in the direction of the earth's magnetic field, which is taken as the z axis. The tensor $\mathbf{T}$ is then given by:

$$\mathbf{T}=-\frac{e}{m\omega^2}\frac{1}{U(U^2-Y^2)}\begin{pmatrix} U^2 & iYU & 0 \\ -iYU & U^2 & 0 \\ 0 & 0 & U^2-Y^2 \end{pmatrix}, \tag{49}$$

(see for example, Banerjea [2]) where $U=1-iZ$, $Z=\nu/\omega$, ν is the average number of collisions made per unit time by an electron, $Y=\omega_H/\omega$, and $\omega_H/2\pi$ is the magnetic gyrofrequency for electrons. This case is of interest in the study of the propagation of whistlers [5].

In a homogeneous medium, a wave whose normal is parallel to the magnetic field is circularly polarized, and the electric field has no component in the direction of the wave normal. Now $\mathbf{E}_1$ is the first approximation to the electric field of the wave and is, therefore, the field of a wave in a homogeneous medium. Hence, it has no component parallel to the z axis, and (27) and (49) show that $\mathbf{r}_1$ similarly has no z component. Hence we may write:

$$\mathbf{r}_1=\mathbf{i}\,x_1+\mathbf{j}\,y_1. \tag{50}$$

There is strong evidence that irregularities in the ionosphere are elongated in the direction of the earth's magnetic field [4]. To allow for this, eq (26) is replaced by

$$\mathbf{K}=u\,\mathbf{i}+v\,\mathbf{j}+f\,w\,\mathbf{k}, \tag{51}$$

where f is a number less than unity. It will be assumed that the irregularities of ΔN are isotropic with respect to the variables u, v, w. The effect of the factor f is as though a region of truly isotropic irregularities had been stretched in the z direction by a factor $1/f$, and f will therefore be called the "elongation factor".

The expression (33) for $\overline{\Delta\mathbf{E}_2}$ must now be written down for this special case. The notation T_{ij} $(i, j=x, y, z)$ will be used for elements of $\mathbf{T}$. Equation (49) shows that $T_{xz}=T_{yz}=T_{zx}=T_{zy}=0$, $T_{xy}=-T_{yx}$. Hence the factor $(\mathbf{T}\,\mathbf{K})\cdot\mathbf{K}$ in (33) becomes

$$(\mathbf{T}\,\mathbf{K})\cdot\mathbf{K}=T_{xx}u^2+T_{yy}v^2+T_{zz}f^2w^2. \tag{52}$$

Equations (27) and (50) show that the factor $\{(\mathbf{T}\,\mathbf{E}_1)\cdot\mathbf{K}\}\cdot\mathbf{K}$ in (33) is given by

$$(x_1u+y_1v)\,(u\mathbf{i}+v\mathbf{j}+fw\mathbf{k}). \tag{53}$$

Now when the summation in (33) is completed, terms involving the products uv, vw, wu will give zero, because ΔN is isotropic with respect to u, v, w. Hence (53) becomes:

$$\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}\cdot\mathbf{K}=x_1u^2\mathbf{i}+y_1v^2\mathbf{j}. \tag{54}$$

Equation (49) shows further that $T_{xx}=T_{yy}$. Hence eq (33) can now be written in full using (52) and (54). It gives

$$\overline{\Delta\mathbf{E}_2}=\frac{e^2}{2\epsilon_0^2}\sum_K\frac{A_K^2}{(u^2+v^2+f^2w^2)^2}\left[\{T_{xx}(u^4+u^2v^2)+T_{zz}f^2u^2w^2\}\mathbf{i}\,x_1+\{T_{xx}(v^4+u^2v^2)+T_{zz}f^2v^2w^2\}\mathbf{j}\,y_1\right]. \tag{55}$$

It is shown in app. B that

$$\sum_K A_K^2\frac{u^4}{(u^2+v^2+f^2w^2)^2}=\sum_K A_K^2\frac{v^4}{(u^2+v^2+f^2w^2)^2}=3\sum_K A_K^2\frac{u^2v^2}{(u^2+v^2+f^2w^2)^2}=\frac{3}{4}L(f)\,\beta N_0^2 \tag{56}$$

$$f^2\sum_K A_K^2\frac{u^2w^2}{(u^2+v^2+f^2w^2)^2}=f^2\sum_K A_K^2\frac{v^2w^2}{(u^2+v^2+f^2w^2)^2}=M(f)\cdot\beta N_0^2 \tag{57}$$

where $L(f)$, $M(f)$ are known functions of f, and β is given by (18). Hence (55) becomes:

$$\overline{\Delta\mathbf{E}_2}=\frac{e^2}{2\epsilon_0^2}\beta N_0^2\{T_{xx}L(f)+T_{zz}M(f)\}\mathbf{r}_1, \tag{58}$$

and (34) (27) give:

$$\mathbf{E}_0=\mathbf{E}_1+\frac{e^2}{2\epsilon_0^2}\beta N_0^2\{T_{xx}L+T_{zz}M\}\mathbf{T}\,\mathbf{E}_1. \tag{59}$$

Next, eq (59) must be written down for this special case. By methods similar to the above it can be shown that

$$\sum_K A_K^2 \frac{\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}\cdot\mathbf{T}\,\mathbf{K}}{\mathbf{K}^2}=\beta N_0^2 Q(f)\{\mathbf{i}(T_{xx}x_1+T_{xy}y_1)+\mathbf{j}(T_{yx}x_1+T_{yy}y_1)\} \tag{60}$$

where

$$\sum_K A_K^2 \frac{u^2}{u^2+v^2+f^2w^2}=\sum_K A_K^2 \frac{v^2}{u^2+v^2+f^2w^2}=Q(f)\cdot\beta N_0^2 \tag{61}$$

(see app B)

Equation (60) is now substituted in (39) and (27), (50) are used. This gives

$$\mathbf{P}_0=eN_0\mathbf{T}\,\mathbf{E}_1-\frac{e^2N_0^2}{2\epsilon_0}\,Q\beta\mathbf{T}^2\,\mathbf{E}_1 \tag{62}$$

which may now be combined with (59) to give the constitutive relation.

The relations (59) and (62) are equivalent to:

$$\mathbf{E}_0=\left\{\mathbf{T}^{-1}+\frac{e^2}{2\epsilon_0^2}\,\beta N_0^2(T_{xx}L+T_{zz}M)\right\}\mathbf{T}\,\mathbf{E}_1, \tag{63}$$

$$\mathbf{P}_0=eN_0\left(1-\frac{eN_0}{2\epsilon_0}\,Q\beta\,\mathbf{T}\right)\mathbf{T}\,\mathbf{E}_1, \tag{64}$$

from which the vector $\mathbf{T}\,\mathbf{E}_1$ is to be eliminated. This then gives

$$\mathbf{E}_0=\mathbf{R}\,\mathbf{P}_0 \tag{65}$$

where $\mathbf{R}$ is to be found. Now β is a small quantity and its square and higher powers will be neglected. Then it can be shown that

$$\mathbf{R}\approx\frac{\mathbf{T}^{-1}}{eN_0}-\frac{\beta}{2\epsilon_0}\left\{X\left(\frac{LU}{U^2-Y^2}+\frac{M}{U}\right)-Q\right\}, \tag{66}$$

and it follows from the expression (49) for $\mathbf{T}$ that

$$\left.\begin{aligned} &R_{xz}=R_{zx}=R_{yz}=R_{zy}=0\\ &R_{xx}=R_{yy};\quad R_{xy}=-R_{yx}\end{aligned}\right\}. \tag{67}$$

The standard methods of magneto-ionic theory may now be used to find the polarization and refractive indices of the waves which can be propagated. The two polarizations are given by $\rho=\pm i$, so that the two waves are circularly polarized with opposite senses, and the two refractive indices are given by

$$\mu^2=1+\frac{1}{\epsilon_0(R_{xx}\pm iR_{xy})}, \tag{68}$$

which leads to

$$\mu^2=1-\frac{X}{U+\frac{1}{2}\beta\left\{X^2\left(\frac{LU}{U^2-Y^2}+\frac{M}{U}\right)-QX\right\}\pm Y}. \tag{69}$$

11. Application to Whistlers

Whistlers are naturally occurring radio signals propagated in the extensive ionized region surrounding the earth, with the wave normals in directions close to the lines of force of the earth's magnetic field [5]. The frequency is very low (often in the audible range) so that X and Y

are very large compared to unity. The waves are not heavily attenuated so that the effect of electron collisions must be small. Hence in (69) we take $U=1$, $X \gg 1$, $Y \gg 1$. Moreover X/Y is inversely proportional to frequency and is very large at low frequencies, so that $X/Y \gg 1$ [5]. One of the values of μ given by (69) is almost purely imaginary; the other value, with $-Y$ in the denominator, refers to the wave which is believed to explain whistlers, and it is given by

$$\mu^2 \approx \frac{X}{Y - \frac{\beta}{2} X^2 \left(M - \frac{L}{Y^2}\right) + \frac{1}{2} \beta Q X}. \tag{70}$$

If the elongation factor f is not near zero, then $M(f)$ and $L(f)$ are both non-zero, but $L/Y^2 \ll M$ since Y is large. X may be neglected in comparison with X^2, so that (70) becomes:

$$\mu^2 \approx \frac{X}{Y - \frac{1}{2}\beta M X^2}. \tag{71}$$

The additional term $\frac{1}{2}\beta M X^2$ in the denominator has an effect similar to that which the Lorentz force would have. Waves could not be propagated if

$$\frac{\beta}{2} M X^2 > Y, \tag{72}$$

for then the refractive index would be purely imaginary. Hence, irregularities could modify the propagation of whistlers. Equation (72) shows that there is a lower limit to the frequencies which could be propagated, given by

$$(\tfrac{1}{2}\beta M f_N^4 / f_H)^{1/3} \tag{73}$$

where f_N is the plasma frequency and f_H is the gyrofrequency.

If, however, the irregularities are very elongated in the direction of the earth's magnetic field, so that the elongation factor f is very small, then $M(f)$ is small (see app. B), and its effect could be opposed by the term L/Y^2 in the denominator of (70). If f is so small that $L \approx MY^2$, then the effect of irregularities on the propagation of whistlers would be small.

It is well known that whistlers are not observed regularly, but only during comparatively short periods of time. They are believed to start from lightning flashes, so that sources of whistlers are always plentiful. The absence of whistlers for long periods must therefore be explained by some feature of the propagation. It is now tentatively suggested that whistlers cannot be observed when the path they would travel contains irregularities which are not greatly elongated. On the occasions when whistlers can be observed, the irregularities are either absent or greatly elongated in the direction of the earth's magnetic field.

12. Conclusion

It has been shown that irregularities in the ionosphere, which are small compared with one wavelength, may modify the constitutive relations and hence may affect the refractive indices for electromagnetic waves. The modifications are in some ways similar to those which would be introduced into the Appleton-Hartree formula by a Lorentz force. When the irregularities extend only in one dimension the theory can be handled by the statistical method of section 5, and it was found that even in a loss-free medium the refractive index now has an imaginary part which might be associated with loss of energy from the wave by scattering. For the general three-dimensional case, the theory is more difficult and has only been handled by a method of successive approximations. A theory similar to that of section 5 is needed for this case but has not so far been found. The method of section 6 onwards is less satisfactory because of the numerous approximations. The results, however, do indicate that small irregularities may play an important part in the propagation of very-low-frequency radio waves in

the ionosphere. In particular, they may explain why whistlers are observed only on comparatively rare occasions.

The author is greatly indebted to the University of Colorado, and to the Directors and Staffs of the High Altitude Observatory and NBS for providing him with the opportunity to work in Boulder.

13. Appendix A: Evaluation of the Integral in Equation (13)

With the notation (14), the integral in (13) becomes

$$\frac{D_0}{a\zeta\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{\Omega+\xi/a\sqrt{2}}\,e^{-\xi^2/2a^2}\cdot d\xi. \tag{A1}$$

Let $\xi/\sqrt{2}a=w$. Then (A1) gives

$$\frac{D_0}{\sqrt{\pi}\,\zeta}\int_{-\infty}^{\infty}\frac{1}{w+\Omega}\,e^{-w^2}\cdot dw. \tag{A2}$$

Now

$$\frac{1}{w+\Omega}=i\int_0^{\pm\infty}e^{-i\tau(w+\Omega)}d\tau, \tag{A3}$$

where the top limit is $+\infty$ when $I(\Omega)$ is negative, and $-\infty$ when $I(\Omega)$ is positive. Here $I(\Omega)$ is always positive. Hence

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{1}{w+\Omega}\,e^{-w^2}dw&=i\int_0^{-\infty}e^{-i\Omega\tau}\left[\int_{-\infty}^{\infty}e^{-w^2-i\tau w}dw\right]d\tau\\&=i\sqrt{\pi}\int_0^{-\infty}e^{(-1/4\tau^2-i\Omega\tau)}\cdot d\tau=2i\pi^{1/2}e^{-\Omega^2}\int_{i\Omega}^{-\infty}e^{-v^2}dv\\&=-\pi^{1/2}[i\pi^{1/2}e^{-\Omega^2}-2F(\Omega)]\end{aligned} \tag{A4}$$

where

$$F(\Omega)=e^{-\Omega^2}\int_0^{\Omega}e^{u^2}\cdot du. \tag{A5}$$

When (A5) is subsitituted in (A2) it leads to the result (15).

The function $F(\Omega)$ has been studied by Mitchell and Zemansky [3] who give the following formulas:
when Ω is small;

$$1-2\Omega F(\Omega)=1-\frac{2\Omega^2}{1}+\frac{(2\Omega^2)^2}{1\cdot3}-\frac{(2\Omega^2)^3}{1\cdot3\cdot5}+\cdots \tag{A6}$$

and when Ω is large;

$$F(\Omega)\sim\frac{1}{2\Omega}\left[1+\frac{1}{(2\Omega^2)}+\frac{1\cdot3}{(2\Omega^2)^2}+\frac{1\cdot3\cdot5}{(2\Omega^2)^3}+\cdots\right]. \tag{A7}$$

14. Appendix B: Evaluation of Coefficents in Equations (55) and (60)

According to eq (51) the vector $\mathbf{K}=\mathbf{i}u+\mathbf{i}v+\mathbf{k}fw$. Its direction cosines are therefore

$$l=\frac{u}{(u^2+v^2+f^2w^2)^{1/2}},\quad m=\frac{v}{(u^2+v^2+f^2w^2)^{1/2}},\quad n=\frac{fw}{(u^2+v^2+f^2w^2)^{1/2}} \tag{B1}$$

These may be written in spherical polar coordinates:

$$l=\sin\theta\cos\phi,\quad m=\sin\theta\sin\phi,\quad n=\cos\theta. \tag{B2}$$

In equation (55) the coefficient $u^4/(u^2+v^2+f^2w^2)^2$ is $l^4=\sin^4\theta\cos^4\phi$. It is assumed that the distribution of the directions of K is the same for all values of A_K, which means that

$$\sum_K A_K^2 l^4=\overline{\sin^4\theta\cos^4\phi}\sum_K A_K^2. \tag{B3}$$

Since all values of ϕ are equally probable this is equal to

$$\sum_K A_K^2 m^4=\overline{\sin^4\theta\sin^4\phi}\sum_K A_K^2. \tag{B4}$$

It can be shown that $\overline{\sin^4\phi}=\overline{\cos^4\phi}=3/8$; $\overline{\cos^2\phi\,\sin^2\phi}=1/8$; $\overline{\cos^2\phi}=\overline{\sin^2\phi}=1/2$. Hence:

$$\sum A_K^2 l^4 = \sum A_K^2 m^4=3/8\,\overline{\sin^4\theta}\sum A_K^2, \tag{B5}$$

$$\sum A_K^2 l^2 m^2=1/8\,\overline{\sin^4\theta}\sum A_K^2, \tag{B6}$$

$$\sum A_K^2 l^2 n^2 = \sum A_K^2 m^2 n^2=1/2\,\overline{\cos^2\theta\sin^2\theta}\sum A_K^2, \tag{B7}$$

$$\sum A_K^2 l^2 = \sum A_K^2 m^2=1/2\,\overline{\sin^2\theta}\sum A_K^2, \tag{B8}$$

$$\sum A_K^2 n^2 = \overline{\cos^2\theta}\sum A_K^2. \tag{B9}$$

Now let

$$fz=z' \tag{B10}$$

so that

$$\mathbf{K}=u\mathbf{i}+v\mathbf{j}+w\mathbf{k}', \tag{B11}$$

where **i, j, k′** are unit vectors parallel to the axes of the coordinate system (x, y, z'). Let the spherical polar angles in this system be θ', ϕ. Then

$$f\tan\theta=\tan\theta'. \tag{B12}$$

Now the irregularities are isotropic with respect of u, v, w, which means that all directions of the vector **K** in the system x, y, z' are equally probable. Hence, the probability $p(\theta', \phi)d\theta' d\phi$ that θ', ϕ, lie in given ranges $d\theta' d\phi$ is given by

$$p(\theta', \phi)=(1/4\pi)\sin\theta'. \tag{B13}$$

The corresponding probability that θ,ϕ lie in ranges $d\theta,d\phi$ is $p(\theta,\phi)\ d\theta d\phi$ where

$$p(\theta,\phi)=p(\theta', \phi)\frac{d\theta'}{d\theta}=\frac{f^2\sin\theta}{4\pi(\cos^2\theta+f^2\sin^2\theta)^{3/2}}. \tag{B14}$$

The average values in eq (B5) to (B9) can now be found. For example

$$\overline{\sin^4\theta}=\int_0^{\pi}\frac{f^2\sin^5\theta}{4\pi(\cos^2\theta+f^2\sin^2\theta)^{3/2}}d\theta\int_0^{2\pi}d\phi. \tag{B15}$$

This reduces to

$$\overline{\sin^4\theta}=L(f)=(1+\tfrac{1}{2}f^2)(1-f^2)^{-2}-2f^2(1-\tfrac{1}{4}f^2)(1-f^2)^{-5/2}\log[\{1+(1-f^2)^{1/2}\}/f] \tag{B16}$$

Similarly it can be shown that

$$\overline{\cos^2\theta\sin^2\theta}=M(f)=-\frac{3}{2}f^2(1-f^2)^{-2}+f^2\left(1+\frac{1}{2}f^2\right)(1-f^2)^{-5/2}\log[\{1+(1-f^2)^{1/2}\}/f] \tag{B17}$$

$$\overline{\cos^4\theta}=P(f)=1-L(f)-M(f) \tag{B18}$$

$$\overline{\sin^4\theta}=Q(f)=L(f)+M(f) \tag{B19}$$

In particular: $L(0)=1$; $L(1)=8/15$; $M(0)=0$; $M(1)=2/15$.
Figure 3 gives curves showing how $L(f)$, $M(f)$ depend upon f.

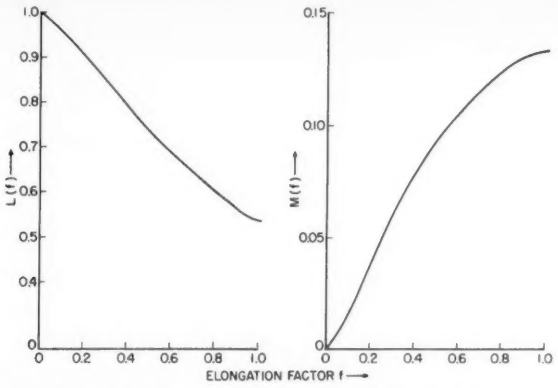


FIGURE 3. *Curves showing how the functions $L(f)$ and $M(f)$ depend on the elongation factor f.*

15. References

[1] E. V. Appleton, Wireless studies of the ionosphere, J. Inst. Elec. Engrs. (London) **71,** 642 (1932).
[2] B. K. Banerjea, On the propagation of electromagnetic waves through the atmosphere, Proc. Roy. Soc. (London) **[A] 190,** 67 (1947).
[3] A. C. G. Mitchell and M. W. Zemansky, Resonance radiation and excited atoms, Cambridge Univ. Press (1934).
[4] M. Spencer, The shape of irregularities in the upper atmosphere, Proc. Phys. Soc. (London) **[B] 68,** 493 (1955).
[5] L. R. C. Storey, An investigation of whistling atmospherics, Phil. Trans. Roy. Soc. London Ser. **[A] 246,** 113 (1953).

BOULDER, COLO.

(Paper 63D2-14)

508098—59——3

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Ionospheric Investigations Using the Sweep-Frequency Pulse Technique at Oblique Incidence

Vaughn Agy and Kenneth Davies

(March 12, 1959)

This paper describes the present state of oblique-incidence investigations of the ionosphere, using the sweep-frequency pulse technique, with special reference to the work carried out at the National Bureau of Standards. After a short review of the published literature, oblique-incidence sweeps are presented showing the diurnal and seasonal variations on two east-west paths of lengths 1,150 kilometers and 2,370 kilometers. The discrepancies between observed and calculated maximum usable frequencies are presented for both paths and then various phenomena of interest are shown. Finally, the above phenomena are discussed in the light of existing knowledge and theory and, in particular, it is shown that the discrepancies between observed and calculated maximum usable frequencies are unlikely to be caused by magnetoionic deviation of the ray.

1. Introduction

The sweep-frequency pulse technique for studying the ionosphere is of vital importance in determining the correct application of vertical-incidence ionospheric data to radio communication problems. For a clear understanding of the important factors which govern the operation of high-frequency radio circuits (signal strength, fading, etc.) direct observations of ionospheric propagation are essential.

With the above problems in mind oblique-incidence observations have been made in several countries. The earliest experiments using the sweep-frequency pulse technique appear to be those of Farmer and Ratcliffe [1] [1] Farmer, Childs, and Cowie [2], and those of Eckersley et al. [3] prior to World War II. During the war Beynon [4] extended these experiments to longer path lengths with improvements in the recording technique. After World War II oblique-incidence work was extended to still greater distances in Germany [5, 6, 7, 8, 9], Canada [10, 11, 12], the United States [13, 14, 15], and Japan [16]. The results of these investigations will be considered in the following section. In the present paper we shall survey the work done in this field with special reference to the experiments carried out by the Bureau.

2. Survey of Previous Work

In the earliest experimental work carried out by Farmer and Ratcliffe [1], pulses were transmitted on an approximately north-south path over a distance of 464 km. Farmer, Childs, and Cowie [2] improved the technique by shortening the sweep time and found that, for ionospherically quiet days, there was good agreement between the observed maximum frequencies for the $F2$-layer and those calculated from a knowledge of the ionosphere overhead at the two ends of the circuit. These observations were confirmed by Eckersley and others [3]. It was found that the average discrepancy between the observed and calculated maximum frequencies was about 0.5 percent. Beynon's observations [4] on a north-south path 715 km long showed that the mean $F2$ maximum usable frequency (MUF) was in excellent agreement with that calculated from end-point vertical-incidence data. Beynon found that the mean of 110 daytime measurements of the MUF factor (MUF/$foF2$) diverged by only 0.2 percent from that calculated from vertical-incidence data using the parabolic layer approximation. In the case of the E-layer maximum frequencies, no close relationship with the vertical-incidence data was found, probably because of sporadic-E reflections. Indeed, weak Es reflections were found to be almost always present.

A more comprehensive series of oblique-incidence observations was carried out in Europe by Dieminger [5, 6, 7, 8, 9] and his coworkers. They made observations over north-south paths of lengths 1,300 km (Helsinki to Lindau) and 1,900 km (Sodankylä to Lindau) and have found that, in general, the agreement between the observed sweeps and those calculated from vertical-incidence soundings made near the midpoint was good except near the junction point of the high-angle and low-angle traces (MUF) [7]. An "inverted" transmission curve was used to derive the ordinary ray critical frequency from the oblique-incidence sweeps. The derived critical frequencies for both the $F1$- and $F2$-layers were about 1 percent higher than those observed. On the other hand, the median observed MUF's were, on the average, about 4 percent higher than the calculated median value. Echo traces extending from the junction point towards higher frequencies were frequently observed for the 1,300-km path and more frequently for the longer path. The top frequencies of these extensions were up to 10 percent higher than the respective MUF's. No correlation was found between these extensions and sporadic E. Dieminger and Möller [6] found that, under certain conditions, the high-angle ray was present

[1] Figures in brackets indicate the literature references at the end of this paper.

over a relatively large frequency range. Furthermore, they showed that during summer midday, at low sunspot activity, the *E*-layer controlled the MUF for a distance of 1,300 km whereas at vertical incidence the highest frequency was reflected from the *F*2-layer.

The Canadian experiments [10, 11] made over an east-west path of 2,350 km showed that under certain conditions the Pedersen ray was relatively important and that signals were sometimes received at frequencies up to 10 percent higher than the so-called "classical MUF". In these experiments, good agreement was found between the observed classical MUF's and the MUF's calculated on the basis of Sellmeyer theory. No departures from reciprocity were observed, and it was shown that the frequency separation between the ordinary and the extraordinary traces at the MUF was about 0.3 Mc.

Records showing the importance of the one-hop Pedersen ray in the determination of the MUF were obtained by Warren and Hagg [12] on a 5,300-km path across the North Atlantic. Over this distance the low-angle ray is cut off by the earth. These observations agreed relatively well with calculations by Kift [17].

In the United States the National Bureau of Standards has conducted oblique-incidence experiments since 1951. These were carried out over two east-west paths; namely, Sterling, Va. (near Washington, D.C.) to St. Louis, Mo., 1,150 km [13, 14] and Sterling, Va. to Boulder, Colo., 2,370 km [15]. The records obtained from these experiments have been scaled primarily for MUF's. Wieder [14] has shown that for transmission over 1,150 km the observed MUF's were sometimes 5 percent higher than the MUF's calculated from midpoint data and the appropriate Smith transmission curve [18].

When the Sterling to Boulder experiment was started, it became apparent that care must be taken to distinguish between two frequencies either of which may, with reason, be called the MUF of a given layer. The first of these, referred to variously as "the junction frequency" (of the high-angle and low-angle rays), or the "nose frequency" (because of the shape of the trace seen in a clear sweep-frequency record), is here called the "classical MUF". It is a quantity determined primarily by the geometry of the ray path between transmitter and receiver and is independent of equipment sensitivity or power. The second of the two frequencies is the highest frequency propagated by a given layer. Terms applied to this frequency include both "maximum observed frequency" (MOF) and "practical MUF" for the layer. For the shorter path it is almost always identical with the classical MUF described above but for the longer path it is frequently higher than the classical MUF by as much as 10 percent. When the latter situation holds, the amount by which the maximum observed frequency exceeds the classical MUF is evidently dependent on the power and sensitivity of the equipment used. The difference between the MOF and the classical MUF (expressed either as frequency difference or percentage frequency difference) is referred to as the "MUF extension" or occasionally as the "nose extension". The appearance of the trace on sweep-frequency oblique-incidence records at frequencies above the classical MUF suggests these terms as well as terms like "whisker" and "tail".

A third quantity of importance is the "calculated MUF" [2] which is the result obtained by application of any one of several methods of computation to the vertical-incidence (midpoint) ionogram (e.g., Appleton-Beynon [19], Smith [20]). When the calculated MUF is derived by means of standard transmission curves of the type devised by Smith, the term "standard MUF" applies. To the extent that theory and experiment agree, the calculated MUF is identical with the classical MUF.

The nose extension was observed only on the longer (2,370-km) path, at least to an appreciable extent. However, Wieder's analysis showed that even in the case of a well-defined MUF (no extension), over the 1,150-km path the average value of the MUF was greater than the value deduced from the vertical-incidence data at the midpoint of the path by between 0 and 10 percent, depending on time of day and season. The mean difference is about 3 percent. On the longer path Sulzer [15] observed echoes on numerous occasions at frequencies considerably above the so-called classical MUF.

The sweep-frequency technique has been used in Japan over distances of about 1,090 km and about 1,840 km, where it has been found that the mean percent differences between the observed and calculated MUF's for *F*2-layer propagation is about 3 percent for the 1,090-km path and about 6 percent for the 1,840-km path. Excellent agreement appears to have been found between the observed and calculated regular *E*-layer MUF's. However, when sporadic *E* was present, appreciable discrepancies were observed.

Summary of results of previous work:

(1) The existing theories of propagation are adequate for calculating MUF's for distances up to about 700 km. For distances of about 1,000 km the MUF errors average approximately 3 percent and appear to increase with increasing path length.

(2) MUF extensions are sometimes observed over distances of the order of 1,000 to 2,000 km. For greater distances such extensions are commonly observed.

(3) The high-angle ray is important under certain conditions.

(4) For certain distances the *E*- and *F*1-layers determine the MUF, especially in summer daytime during periods of low sunspot activity.

(5) The frequency separation between the ordinary and extraordinary rays at the MUF's decreases with increase of distance from about 0.7 Mc at vertical incidence to about 0.3 Mc at 2,400 km.

[2] The IXth Plenary Assembly of the CCIR (Los Angeles, April 1959) adopted a recommendation on the meaning of MUF, advocating the use of the term "classical MUF" to designate the highest frequency transmitted by ionospheric refraction alone. (Our use of the term is consistent with this definition as applied to a particular layer). The recommendation also suggests that "experimental MUF" be used to designate the results of special experiments (our MOF), and that "theoretical MUF" could be used for the results of theoretical calculations (our "calculated MUF").

3. NBS Oblique-Incidence Program

The Bureau's oblique-incidence experiment has been in operation since August 1951. From August 1951 to December 1952 observations were made over the 1,150-km path between Sterling, Va., and St. Louis, Mo. A conventional vertical-incidence ionosonde was operated at Batavia, Ohio, the midpoint of the great-circle path. From 1953 onwards, the circuit used was between Sterling, Va., and Boulder, Colo., (2,370 km) with the midpoint ionosonde located at Carthage, Ill.

The following features will be considered in the present section: (1) Sample records indicating diurnal and seasonal changes for the two paths; (2) the quantitative relationship between oblique-incidence and vertical-incidence data, including sweep-frequency backscatter data; (3) some interesting phenomena.

3.1. Sample Sweeps

In figures 1 and 2 we see the diurnal variations on quiet summer and quiet winter days for the 1,150-km path. It is seen that near noon in summer the MUF is determined by the *E*-layer and not by the *F*2-layer as is the case at vertical incidence. It should

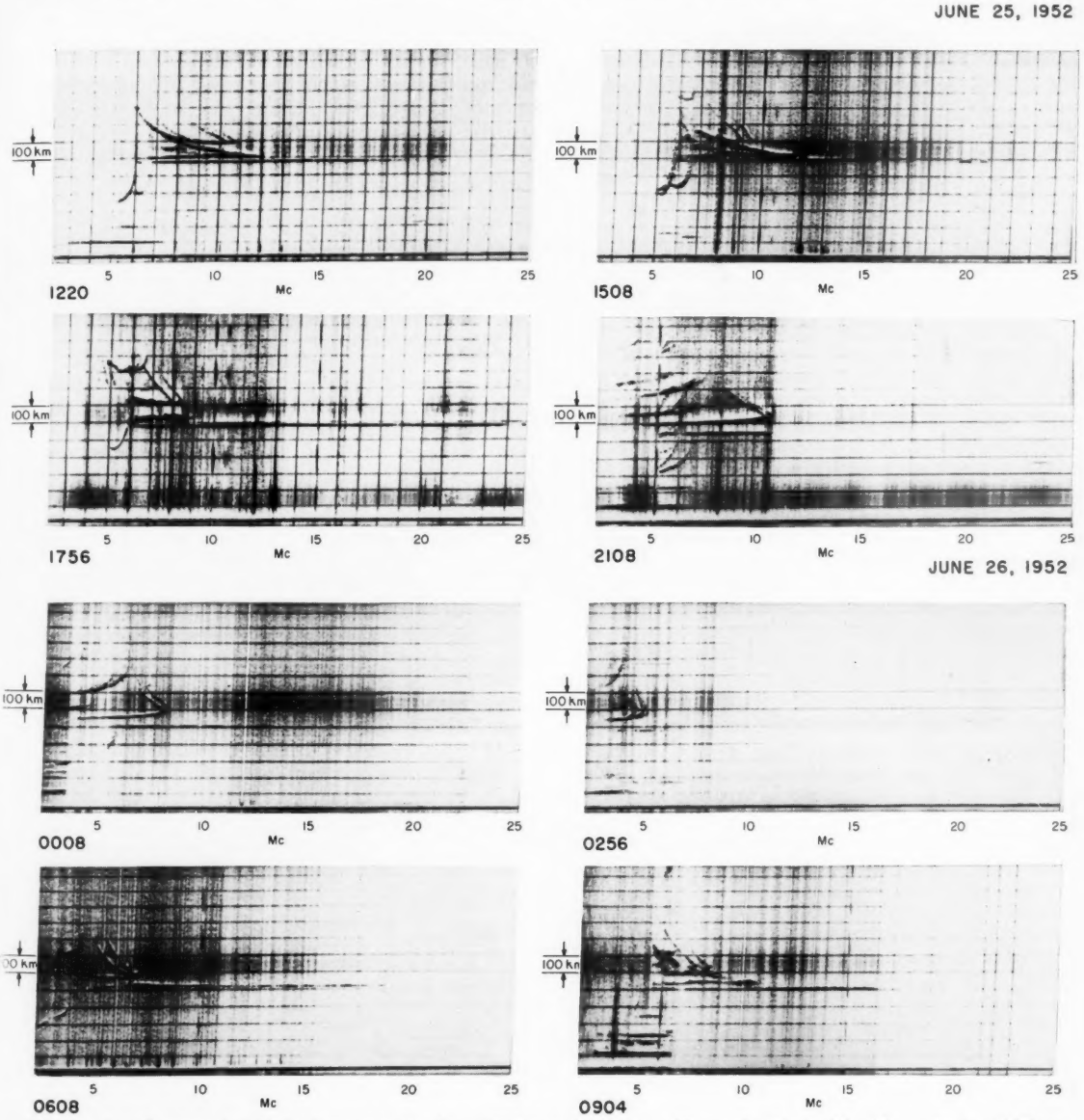


FIGURE 1. *Summer day diurnal sequence for the 1,150-km path, Sterling, Va., to St. Louis, Mo. (June 25, 26, 1952).* The virtual path scale, P', is $\frac{1}{2}(\Delta c t)$ in all sweeps. Times refer to 75° West Meridian Time.

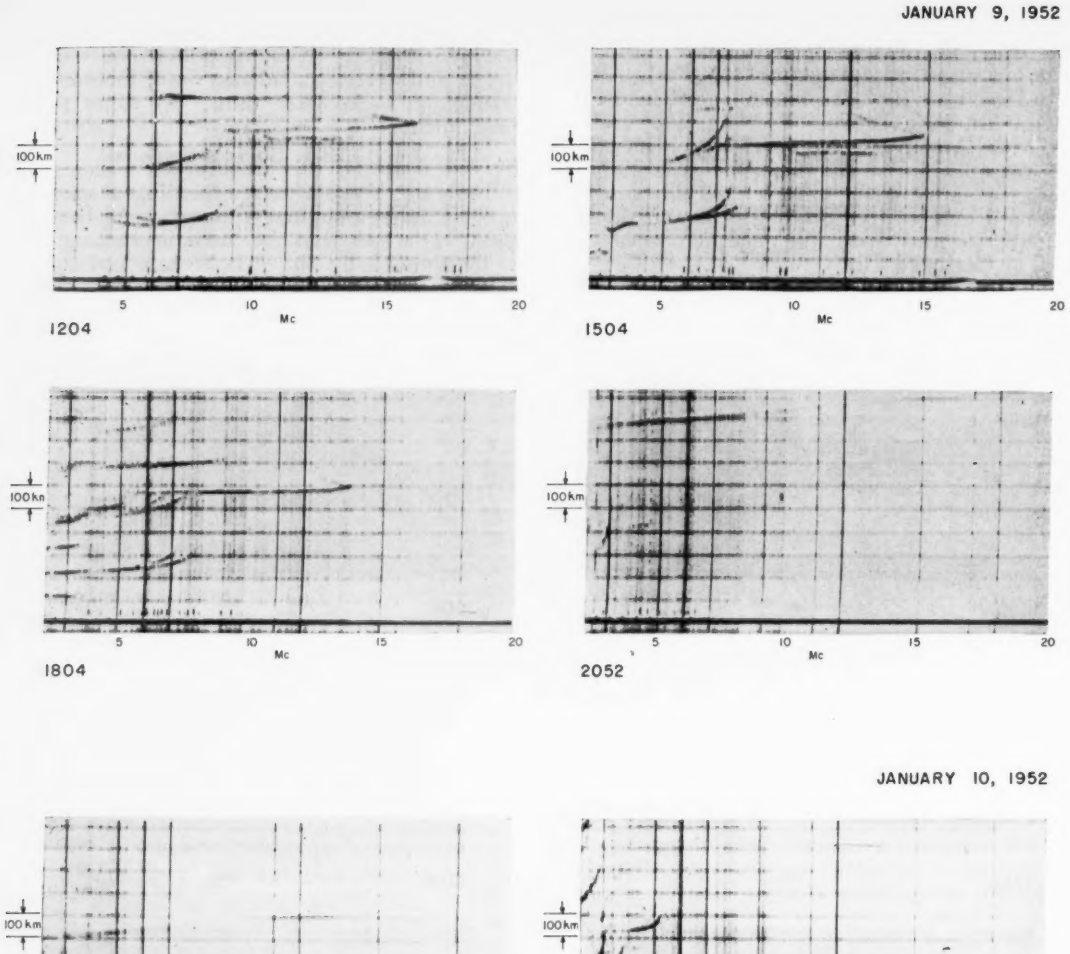


FIGURE 2. *Winter day diurnal sequence for the 1,150-km path, Sterling, Va., to St. Louis, Mo. (Jan. 9, 10, 1952).*

be borne in mind that these observations were made during a period of relatively low sunspot activity. Notice that the $F2$-layer echo is retarded in the underlying $F1$-layer. It will be seen that near noon the $F1$-layer Pedersen ray is relatively strong down to about 6 Mc at which frequency there is little trace of the low-angle ray, probably because of "cut-off" by Es (see sec. 3.3.a). Near sunset the $F1$-layer trace disappears, leaving only the $F2$-layer trace which persists throughout the night. The $F1$-layer is reformed near sunrise. Sometimes the $F1$-layer is continuous with the night $F2$-layer, in these cases the daytime $F2$-layer trace first appears high up on the $F1$-layer Pedersen ray. On other occasions the $F2$-layer is continuous with the night $F2$-layer, and under these conditions the $F1$-layer trace develops below the low-angle $F2$ trace. This appears to have been the case on the morning of June 26, 1952. On a typical winter day (fig. 2) there is little or no evidence of a separate $F1$ trace. It will also be seen that the Pedersen ray is relatively weak as compared with that for a summer day. This is to be expected on the basis of increased defocusing of the high-angle ray for the thinner wintertime layer. Furthermore, the maximum frequency of the $F2$-layer trace occurs near noon in contrast to the case for a summer day, when the highest value is reached near 1800 hr. It is seen that on a winter night the $F2$-layer MUF falls to a low value.

Figures 3 and 4 are similar sequences indicating diurnal variations over the 2,370-km path for a summer day and a winter day. Note that during summer midday, say 0900 to 1500, the $F1$ layer determines the MUF for the path. Here again, it should be remembered that the period under consideration was one of relatively low sunspot activity. Two-hop Es traces are evident (2,370 km is beyond the one-hop limit for reflection heights less than about 110 km); and occasionally the two-hop Es mode may provide the maximum frequency for the path, particularly during summer daylight hours. Also worth noting is the relative difficulty of specifying precisely the frequency at which the upper and lower $F1$ traces meet (the classical MUF). The increasing path length with frequency indicated in the first record of the sequence is caused by a drift of the pulse repetition rate of the Sterling equipment relative to that at Boulder; the corresponding record made at Sterling shows a trace with a downward slope. The average ordinate at a given frequency will quite accurately give a measure of the virtual path (delay time) at that frequency if the reciprocity law is valid.

The winter day sequence (fig. 4) shows little if any $F1$ and the presence of Es is indicated by the M or N trace lying between those representing the one-hop and two-hop $F2$ modes, especially on the record for 0603 90th West Meridian Time.

The records for the longer path (Sterling to Boulder) are often more complicated than those for the shorter path. This additional complication shows itself in the number of traces and the sometimes confusing relationship among them. More important, from the standpoint of MUF data analysis, is the difficulty of accurate scaling of the MUF produced by the fact that the angle between the upper-ray and lower-ray traces becomes more acute as the path length increases. The finite pulse length combined with the more acute angle may, for the $F2$ layer, give an uncertainty in the classical MUF of several tenths of a megacycle and for the $F1$-layer MUF an uncertainty of 1 Mc or more. In those records where it is impossible to define the frequency at which the upper-ray and lower-ray traces meet, the procedure used has been to extrapolate the upper trace along its lower edge (corresponding to the leading edge of the pulse producing it) to intersect with the lower edge of the low-angle trace. Such a procedure provides, if not always accurate values of the classical MUF, at least a degree of consistency which would otherwise be lacking.

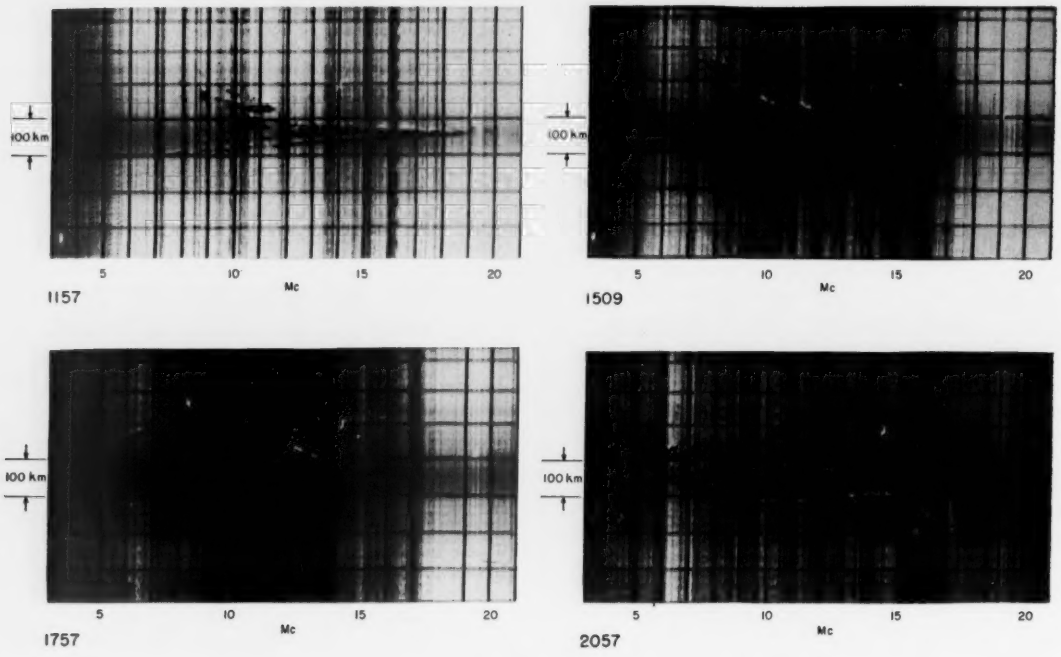
3.2. Vertical and Oblique Incidence Data

The calculation of MUF's is still one of the most important problems in high-frequency radio communication. Because of its simplicity, the Smith transmission curve [18] has been the practical link between vertical-incidence ionospheric data and oblique-incidence predictions in the Central Radio Propagation Laboratory and many other laboratories. The oblique-incidence experiments described herein have afforded a valuable, although limited, check upon the accuracy of results obtained by its use. In the following paragraphs the discrepancies between the MUF's obtained by the use of the Smith transmission curves and the observed MUF's are considered.

The Smith curves may be defined by the equation $f=kf'\sec\Phi$ where f is the oblique-incidence frequency, f' is the equivalent vertical-incidence frequency, Φ is the half-angle at the apex of the equivalent path, and k is a correction factor assumed to depend only on transmission distance.

a. Sterling to St. Louis

Sample winter and summer day plots showing the diurnal variation of observed and calculated MUF are shown in figures 5 and 6. (All times quoted for the Sterling to St. Louis path refer to the 75° West Meridian.) It will be seen that in both cases the measured MUF's are in general higher than those calculated from midpoint records. The average difference between the two MUF's is about 3 percent. It appears that, on the whole, the percent error in the calculated values is greater during the day than during the night. The average daytime monthly error is greater during the winter than during the summer. This is illustrated by the histograms of figure 7. The average error is 3.6 percent for February 1952 and 1.5 percent for May 1952. However, there are many more negative errors in summer so that the spread is much greater.



AUGUST 4-5, 1954

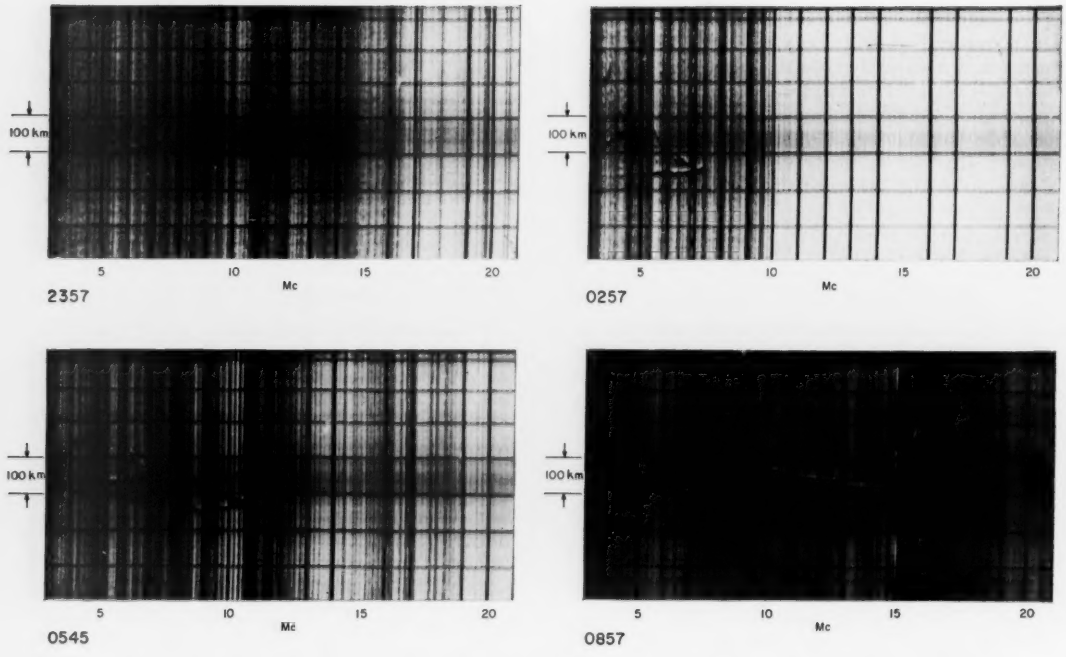
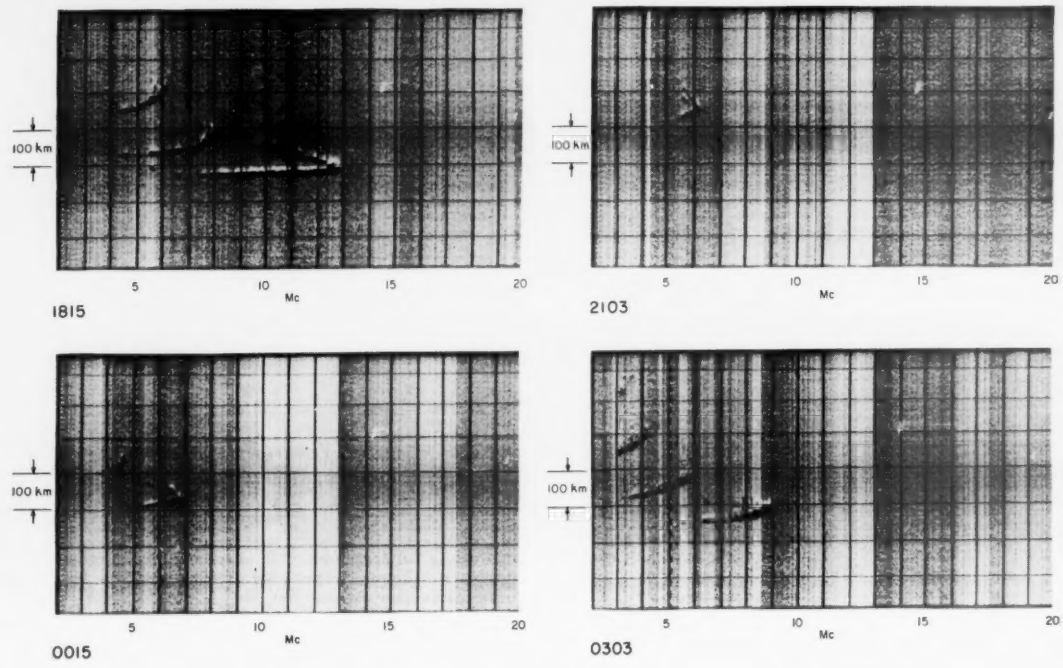


FIGURE 3. *Summer day diurnal sequence for the 2,370-km path, Sterling, Va., to Boulder, Colo. (Aug. 4, 5, 1954).*
Times refer to 90° West Meridian Time.



FEBRUARY 4, 1954

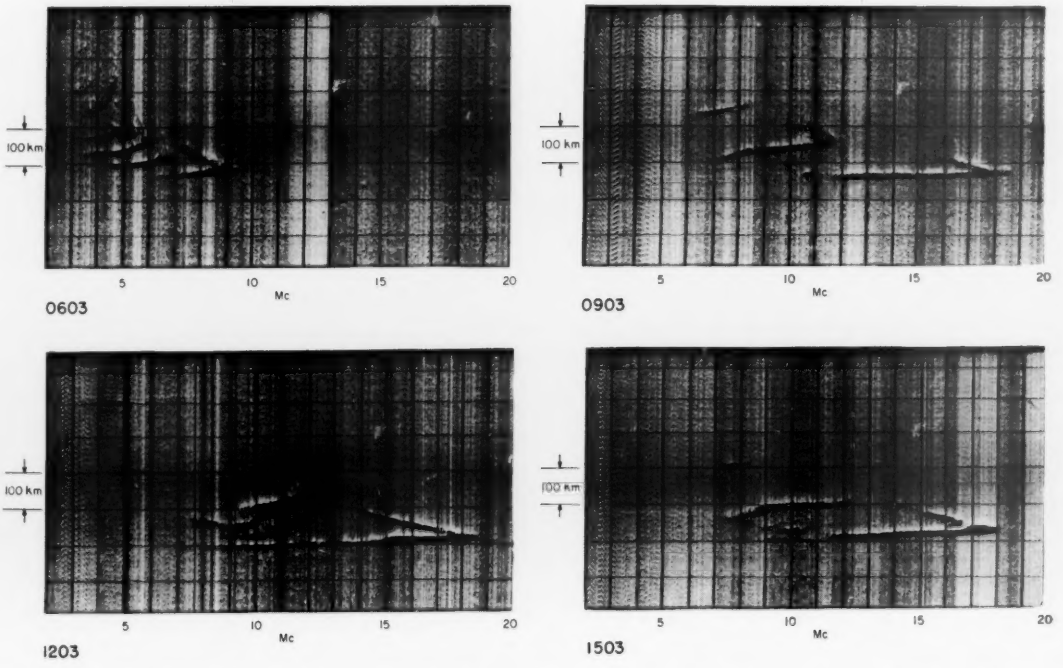


FIGURE 4. *Winter day diurnal sequence for the 2,370-km path, Sterling, Va., to Boulder, Colo. (Feb. 3, 4, 1954).*

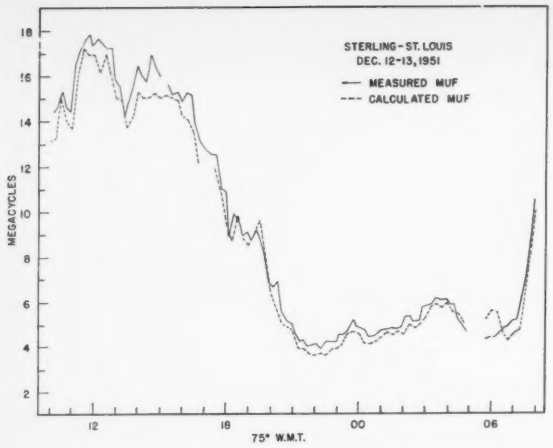


FIGURE 5. *Comparison between the calculated MUF's and the observed MUF's for the 1,150-km path (winter day).*

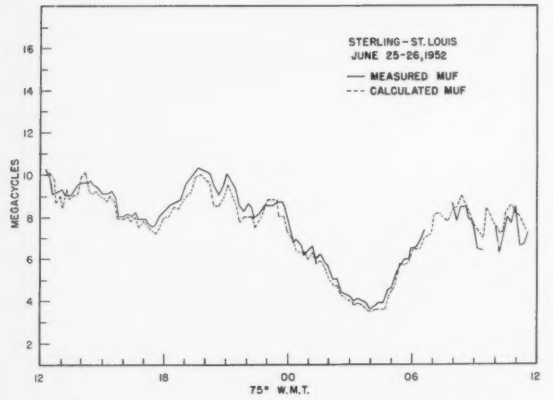


FIGURE 6. *Comparison between the calculated MUF's and the observed MUF's for the 1,150-km path (summer day).*

It will be seen that $F1$-layer MUF's are not included in figures 5 and 6. These data were excluded because of the uncertainty involved in the determination of the MUF for this layer due to the overlapping of the $F1$ trace and the two-hop E trace in the vicinity of the $F1$-layer MUF.

b. Sterling to Boulder

Figures 8 and 9 are sample plots of the observed and calculated $F2$-layer MUF's for the 2,370-km path. The magnitude of MUF extensions when present are indicated by the lengths of short vertical lines drawn upward from the plotted points representing the observed classical MUF's. A comparison of these graphs shows that whereas for a winter day the calculated MUF values are almost always less than the observed values, the same is not necessarily true for a summer day. Indeed, near noon in summer the observed MUF may fall well below the calculated value; this is similar to the observations made on the shorter path. It should be noted that at these times the $F2$-layer echoes are usually heavily retarded in the $F1$ layer, and the virtual height of reflection is thereby increased. It is likely, as indicated in section 3.4, that the correction factor, k, should properly decrease with increasing virtual height. If this adjustment be assumed, the winter to summer differences seen in a comparison of figures 8 and 9 can readily be accounted for. (All times quoted for the Sterling to Boulder path refer to the 90° West Meridian.)

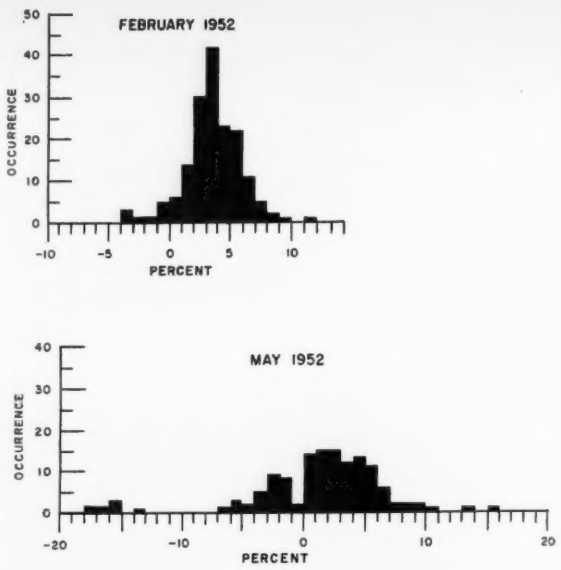


FIGURE 7. *Histograms of 2 month's distributions of the percentage difference between the daytime observed MUF's minus the calculated values for the 1,150-km path, Sterling to St. Louis.*

The histogram in figure 10 presents data from the Sterling to Boulder path. Here, however, the maximum observed frequency for the $F2$ layer (rather than the classical MUF) has been compared with the calculated MUF. Similarity is seen between figure 10 and figure 7 (Sterling to St. Louis), in that the winter day plot shows somewhat less dispersion than that for the summer day. Whereas, the percentages for the summer day are comparable for the two paths, for the winter day the median percent difference for the longer path is about 7 percent compared with 3 percent for the shorter path. However, the difference in the method of derivation (classical MUF for Sterling to St. Louis, MOF for Sterling to Boulder) appears to be largely responsible for the difference shown (see fig. 18). The agreement between the summer-day plots for the two paths is explained by the fact that the "MUF extension" seldom appears on the $F2$-layer trace during summer midday even over the longer path.

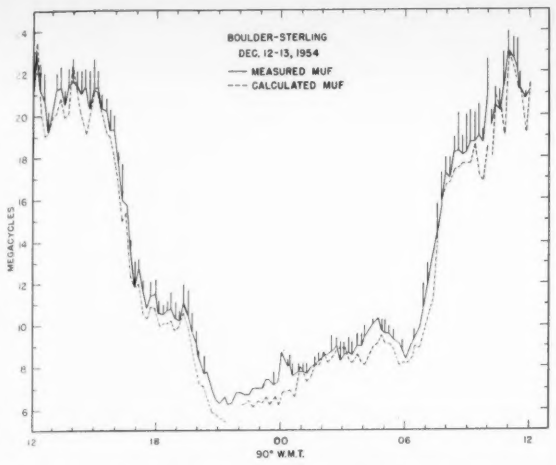


FIGURE 8. *Comparison between the calculated MUF's and the observed MUF's for the 2,370-km path (winter day).*
Vertical lines represent MUF extension.

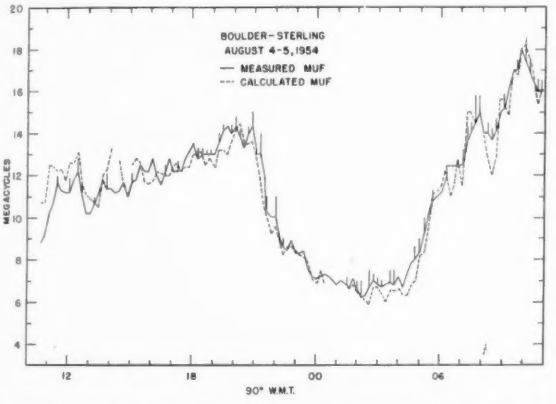


FIGURE 9. *Comparison between the calculated MUF's and the observed MUF's for the 2,370-km path (summer day).*

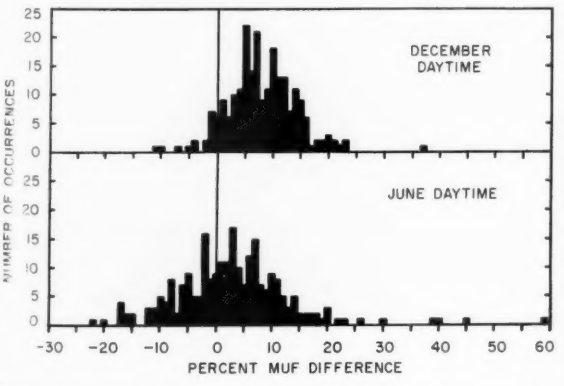


FIGURE 10. *Distribution of percentage difference between the maximum received F2-frequency observed for the Boulder to Sterling path minus the ordinary wave F2 MUF scaled from midpoint records, 1954.*

Another interesting point is brought out in figures 11a and 11b. Here the vertical-incidence ionogram (fig. 11b) would probably be scaled as a blackout (*B*) if it were scaled according to the internationally agreed scaling rules [19]. However, there is a definite echo on the oblique-incidence ionogram. This observation shows that care is needed in applying blackout studies at vertical incidence to communications problems.

c. MUF's Deduced from Sweep-Frequency Backscatter Records

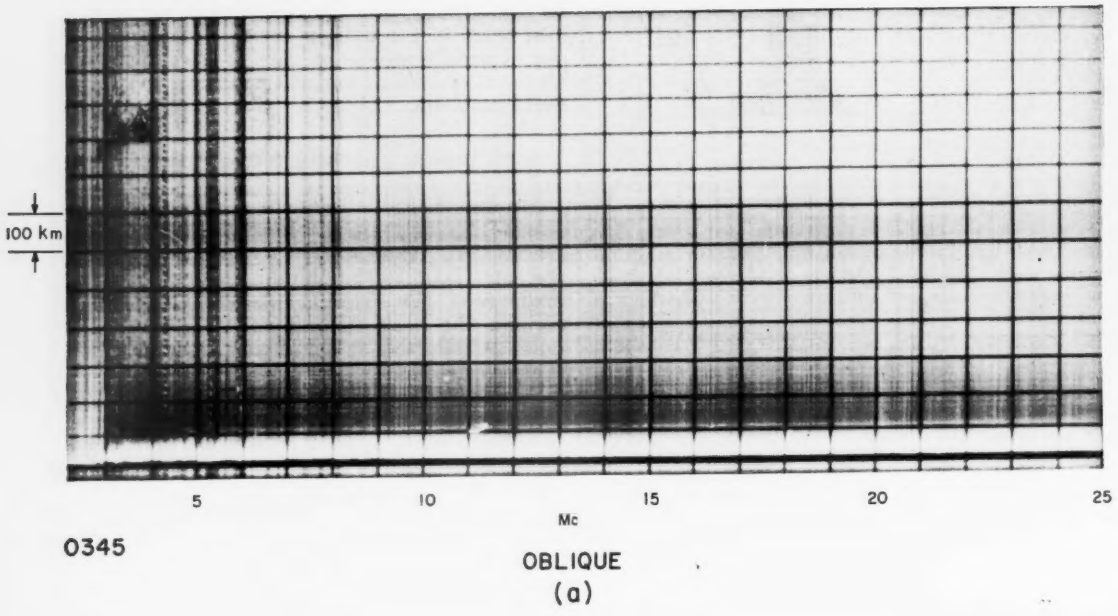
Maximum usable frequencies can be measured by sweep-frequency pulse observations of backscatter echoes [22]. Backscatter recordings were made at the same time as the oblique-incidence records were made over the Sterling to Boulder path. If we assume that the leading edge of the backscatter pulse is reflected at the edge of the skip distance, then the frequency is the MUF for that distance.

Such a backscatter record is shown in figure 12a along with the $P'f$ sweep, figure 12b, for the Boulder to Sterling path. It will be seen that for an assumed slant range of 2,500 km the backscatter record yields an MUF of 16.3 Mc, which is in good agreement with the value of 16.2 Mc taken from the $P'f$ sweep and is slightly larger than the value of 16.0 Mc calculated from the midpoint vertical-incidence data. This result agrees with the observations of Silberstein [23] who found that over this same path the MUF's derived from backscatter data are, on the whole, 0.5 percent greater than those derived from oblique-incidence data. On the other hand, the percent difference between MUF's obtained from vertical-incidence data and those observed at oblique incidence is about −2.0 percent.

d. Spread Echoes

One of the problems facing the frequency predictor is that of determining the MUF from a spread vertical-incidence record. This problem is of particular importance at high latitudes where even the slightest ionospheric disturbance gives rise to spread echoes. The *F*2-layer critical frequency normally scaled is generally near the lower frequency limit of the spread *F*2 trace; and so the question arises whether the normal prediction methods, based on monthly median $foF2$ values, are adequate under these conditions.

At the latitudes over which the Bureau experiments were carried out the occurrence of spread echoes is relatively rare and hence a thorough study of this problem has not been possible. However, one example is shown in figures 13a and 13b, which shows the oblique-incidence ionogram and the corresponding midpoint vertical-incidence ionogram for the 1,150-km path. Although the oblique-incidence traces show signs of spread, the ordinary ray can be identified from the high-angle ray and the classical MUF is seen to be about 4.1 Mc. The vertical-incidence ionogram, on the other hand, shows considerable confusion and it is difficult to decide what



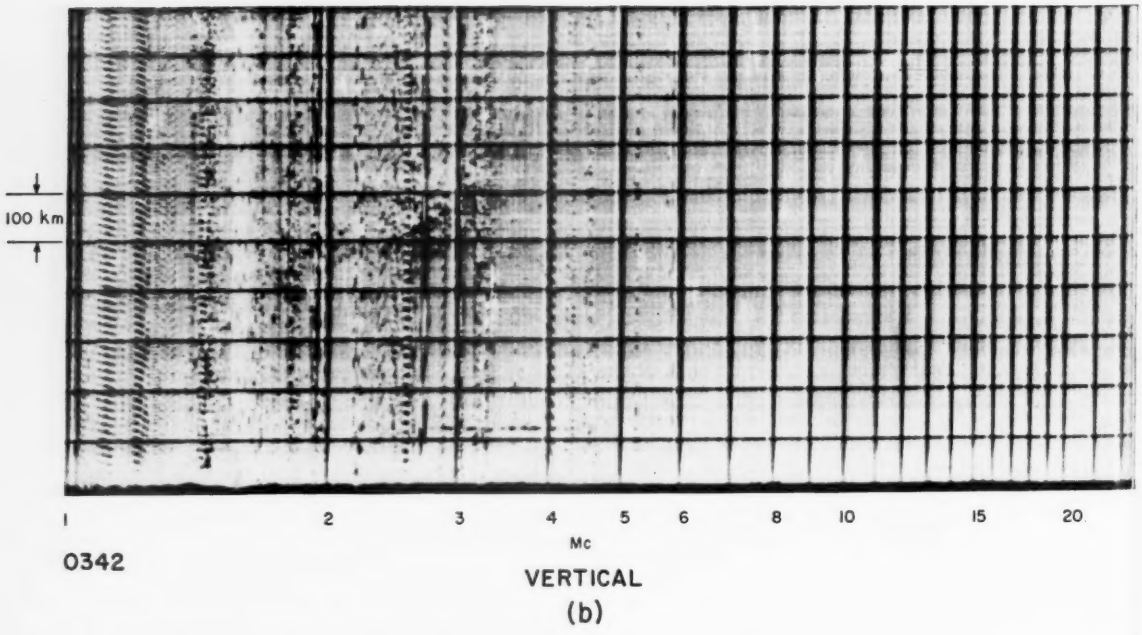
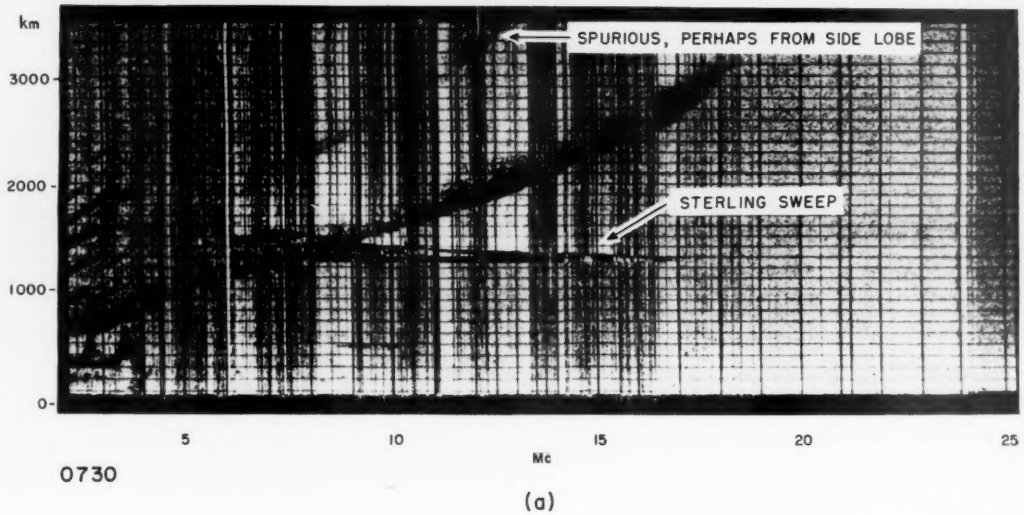


FIGURE 11. *Sweeps showing the presence of echoes on oblique-incidence ionograms (a) and absence on the midpoint vertical-incidence ionograms (b).*

Note that the times shown on the oblique-incidence ionograms refer to the end of the sweep which takes 12 min. whereas the vertical-incidence sweep takes 15 sec.



FEBRUARY 18, 1955

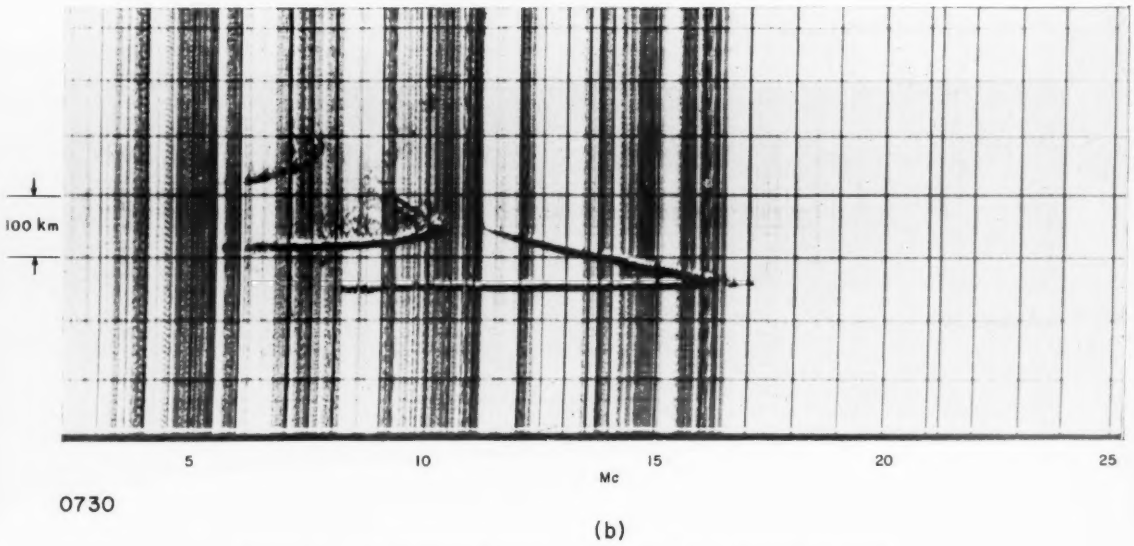
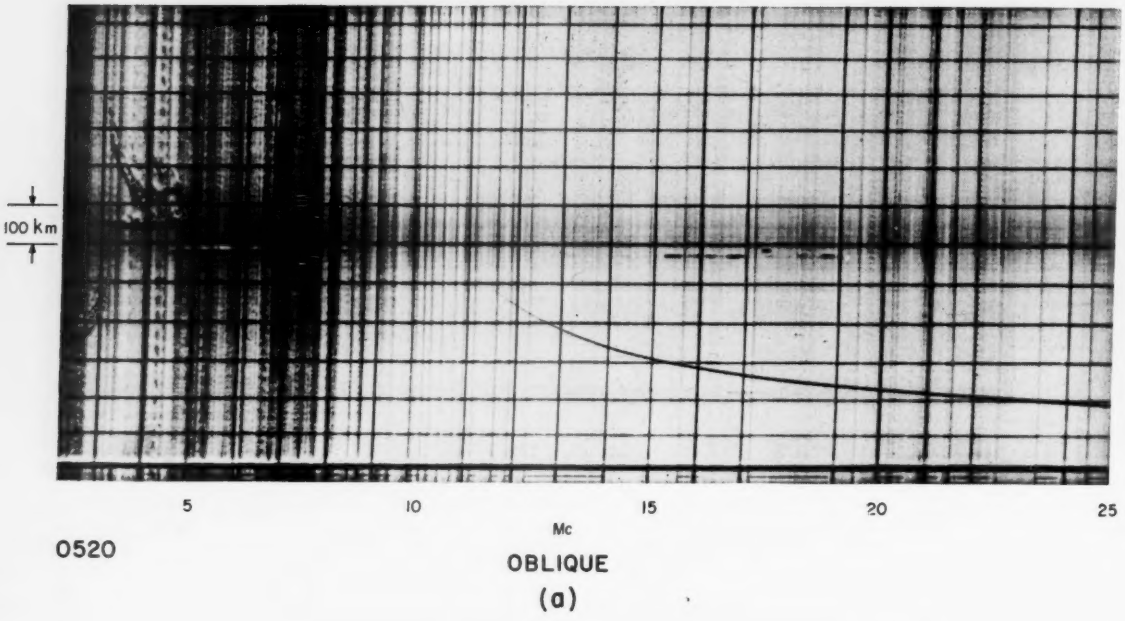


FIGURE 12. (*a*) *Backscatter record.* (*b*) *Boulder to Sterling P'f sweep.*

the critical frequencies are and where to place a transmission curve in order to derive the oblique-incidence MUF. A transmission curve is shown which, when placed on the vertical-incidence ionogram, gives a 1,150-km MUF of 4.1 Mc. It appears that the transmission curve almost touches the inside edge of the spread echo rather than the high-frequency edge. It is of interest to note that there is no feature on the vertical-incidence ionogram giving a frequency of 5.1 Mc, which is the maximum frequency reflected from the *F*2 layer at oblique incidence. Another feature of interest in figure 13a is the presence of sporadic-*E* echoes on the oblique-incidence ionogram with no sign of echoes from the same layer at vertical incidence. A calculation of the separation between the *E*s trace and the *F*2 trace indicates that the lower trace is, in fact, the one-hop *E*s trace.

A similar pair of records for the longer path is shown in figure 14. The transmission curves for



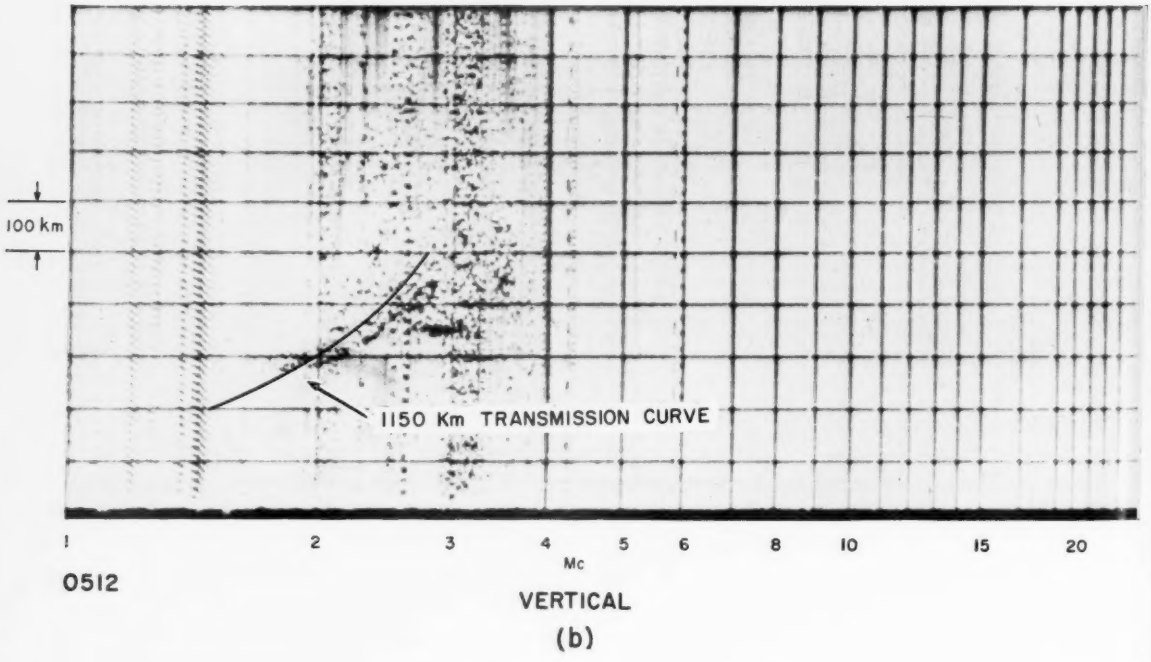
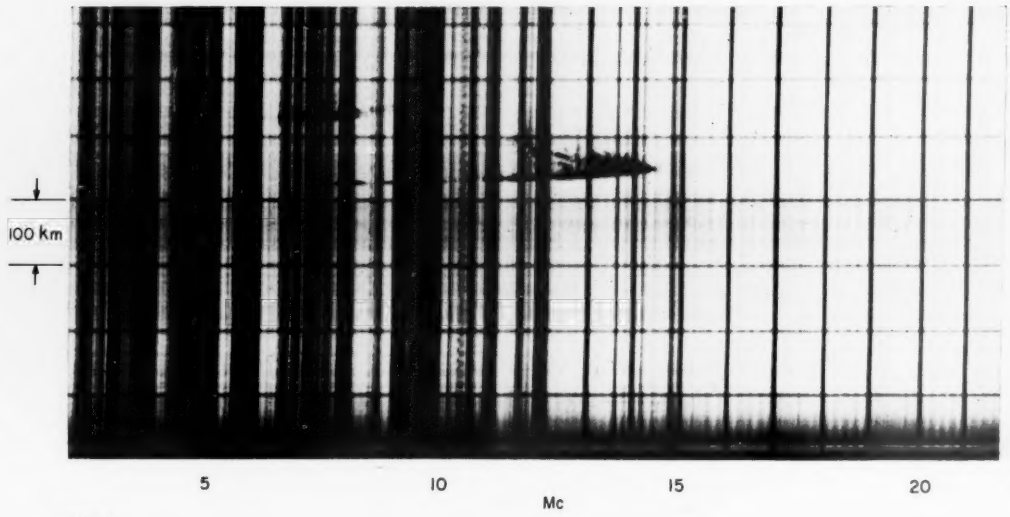


FIGURE 13. *Scaling of the MUF in the presence of spread echo, Sterling to St. Louis.*

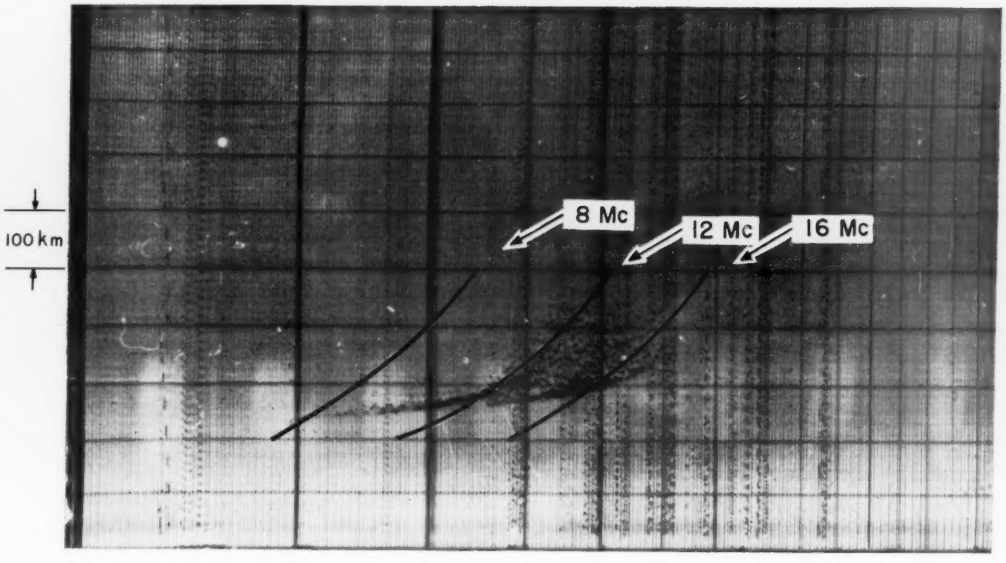
Note the proximity of the 4.1-Mc transmission curve to the inside edge of the spread echo.



2120

OBLIQUE

(a)



2115

VERTICAL

(b)

FIGURE 14. *Scaling of the MUF in the presence of spread echo, Sterling to Boulder.*

8, 12, and 16 Mc have been drawn across the midpoint record. The difficulties of scaling both of the records are apparent: it is not possible to match corresponding features. It should be made clear, however, that the discrepancy between the observed $F2$ MUF of about 14 Mc and the calculated MUF of 12 Mc (inside edge of spread) is no larger than occasionally arises when spread echo is not present.

e. MUF Extensions

It was pointed out that on the longer transmission distances echoes have been observed on frequencies higher than that at which the high-angle and low-angle rays join (the so-called classical MUF). Such extensions have been observed on both the 1,150-km and the 2,370-km paths in the United States.

In figure 15 are sample records taken over the 1,150-km path showing two occurrences of MUF extensions, one on the $F1$-layer trace by day and the other on the $F2$-layer trace in the early morning. In the sweep of 1146, April 3, 1952, we see that the junction of the high-angle and the low-angle rays cannot easily be determined but probably occurs at about 9.5 Mc whereas the maximum observed frequency for the $F1$ trace is near 10.3 Mc. In this case there is difficulty in distinguishing between the $F1$ trace and the two-hop E trace. Note, however, the absence of any such extension on the ordinary- and the extraordinary-$F2$ traces.

Sweeps showing MUF extensions on the 2,370-km path are shown in figure 16. It will be seen that the extension has been observed both by day and by night. Figure 16a is a fairly typical winter-day sweep. We see that the high-angle and low-angle rays merge near 21.1 Mc whereas the maximum observed frequency is about 23.4 Mc. Even if we extrapolate the lower edge of the high-angle ray trace (about 22.0 Mc) there is still an appreciable extension of about 1.4 Mc. Figure 16b is an early morning record which shows an MUF extension about 0.6 and 1.0 Mc depending on the procedure used in determining the MUF (see sec. 3a). It should be emphasized at this point that this phenomenon is not always observed. Figure 17 shows a relatively common sweep taken on the Boulder to Sterling circuit which shows no extension whatsoever.

The extension is greatest during winter days and is only occasionally observed during the middle of the winter nights. The diurnal variation of the median extension of the $F2$ trace for the winter of 1954 to 1955 is shown in figure 18. The extension is practically nonexistent between the hours of 2030 and 0230, and there is a marked dip in the curve near noon. It may be relevant to note that the dip, which occurs between about 0800 and 1600, is coincident with the solar enhancement of the D and E layer ionization. The extension may thus be influenced by the electron distribution below the $F2$ layer.

The histogram in figure 10 for the longer path gives a median winter-day difference between maximum observed frequency and calculated MUF of about 7 percent. It has been pointed out above that whereas this median difference was higher than that indicated in figure 7 for the shorter path, figure 7 (for the Sterling to St. Louis path) is a plot of percentage difference between classical and calculated MUF's. Figure 18 is an indication of the reason for the major difference between figure 7 and figure 10. The 4 percent to 5 percent median MUF extension from figure 18 may be applied as an approximate correction to the 7 percent difference seen in figure 10. The corrected difference of about 3 percent agrees well with the short-path result given by Wieder. For the paths used in the Bureau experiments, it appears then that the error, i.e., the difference between the calculated MUF and the classical MUF, is independent of path length.

3.3. Other Phenomena

a. High-Angle Ray

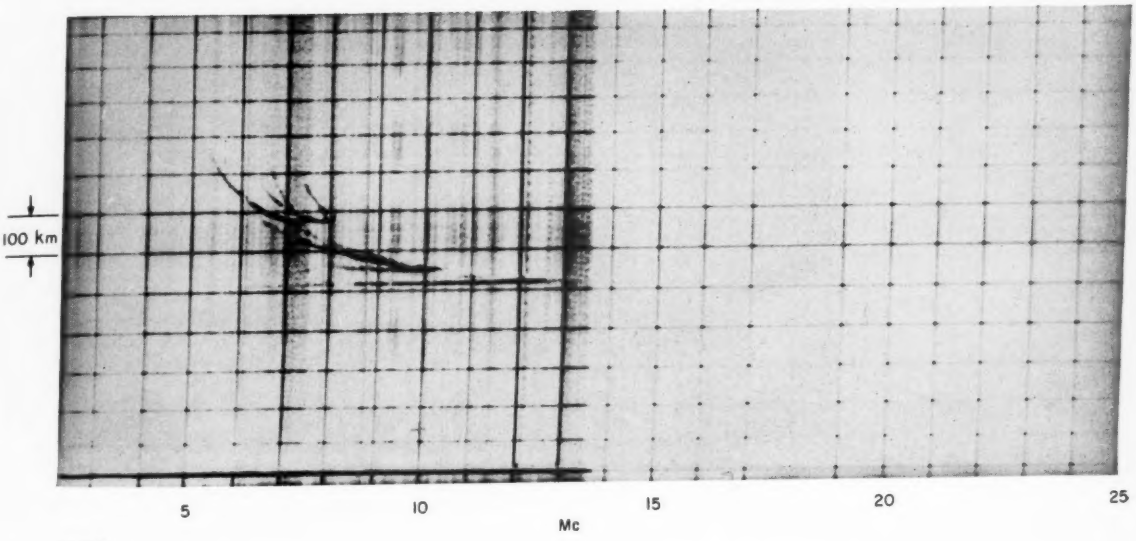
For many years in radio communications forecasting the high-angle ray was neglected in comparison with the low-angle ray on the basis that, except near the maximum frequency, deviative absorption and purely geometrical defocusing would be excessive. This argument is supported by sweeps such as that shown in figure 19 in which only the low-angle ray of the one-hop echo is observed. On the other hand, there are times when the high-angle ray may be important relative to the low-angle ray. Such conditions are especially prevalent during summer days as seen in figures 20a and 20b. In figure 20a we see that for 1,150 km the low-angle ray is completely absent for the one-hop mode, whereas the high-angle ray appears to be relatively strong and decreasing in delay time up to about 9.3 Mc; above that frequency the signal is cut off because of the increase in the angle of incidence. In figure 20b we see that in the frequency range between 5 and 6 Mc the high-angle ray (ordinary wave) appears to be much stronger than the low-angle ray, probably because of reflection of the low-angle signal by the intense sporadic-E which can be seen at frequencies up to about 18 Mc.

This explanation for absence of the low-angle ray is illustrated in figure 21. The low-angle ray is incident on the sporadic-E layer at a larger angle than the high-angle ray and may, therefore, be reflected by Es at substantially higher frequencies than the more steeply incident high-angle ray and will not, under these conditions, be received.

b. Low-Frequency Cut-Off

As the transmitted frequency is increased from its lowest value, we often notice that the first signals to appear at the receiver are those which have suffered two or more reflections at the ionosphere. Two cases of this phenomenon are shown in figures 22 and 23. The former is a case where over the 1,150-km path sporadic-E cuts off the one-hop signal almost completely, whereas the two-hop echo, which penetrates the E region at a different locality and with a smaller angle of incidence, comes in strongly above

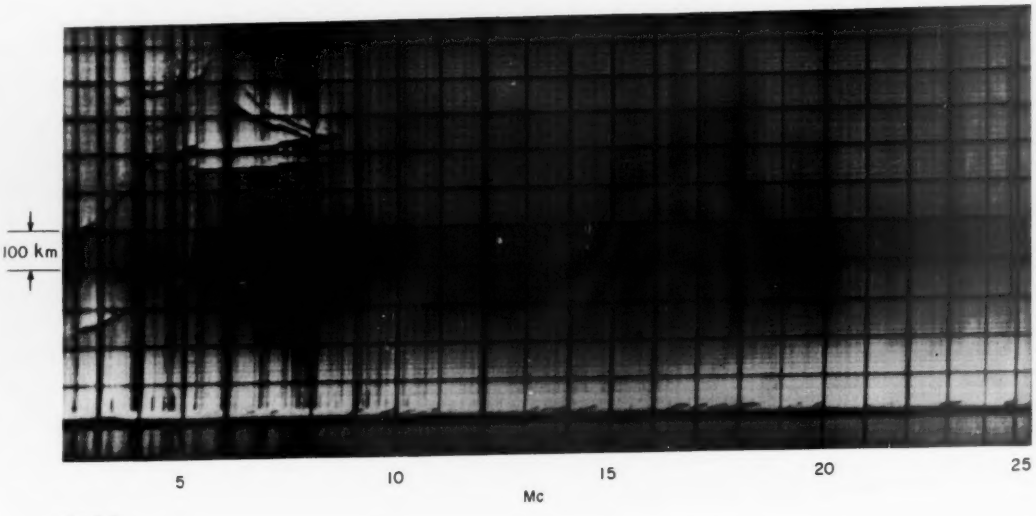
APRIL 3, 1952



1146

(a)

SEPTEMBER 12, 1952

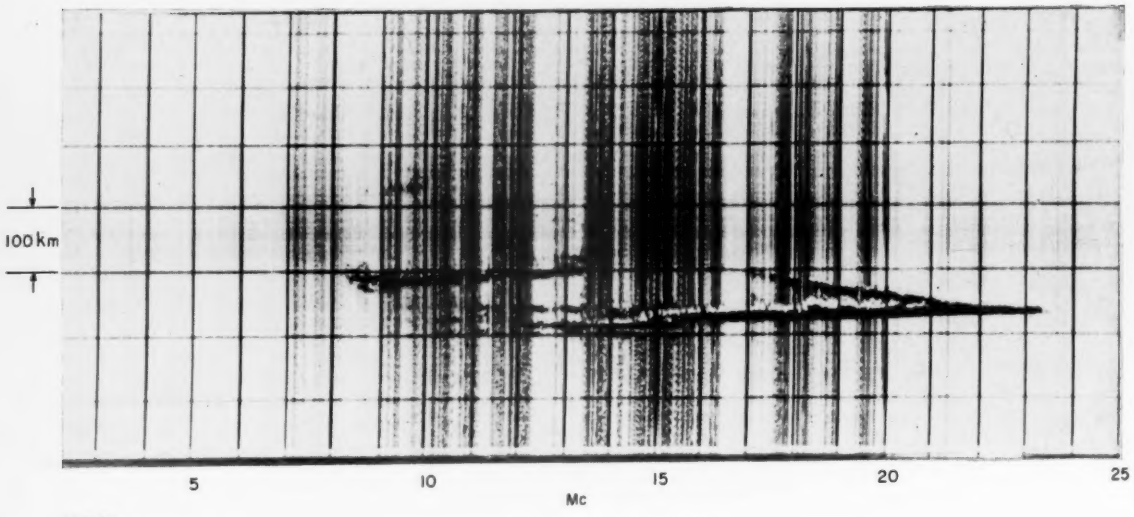


0652

(b)

FIGURE 15. *Sterling to St. Louis sweeps showing the presence of MUF extensions.*

DECEMBER 7, 1954



OCTOBER 5, 1954

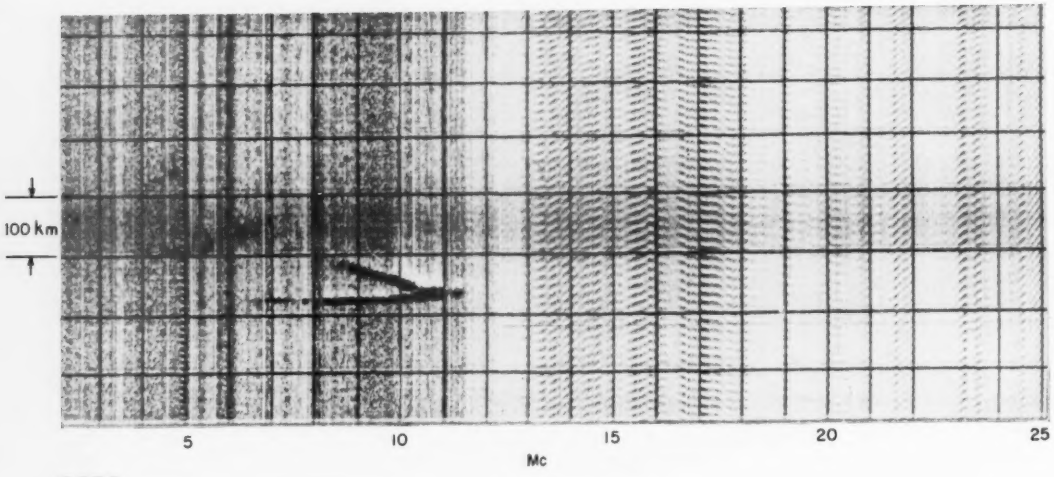
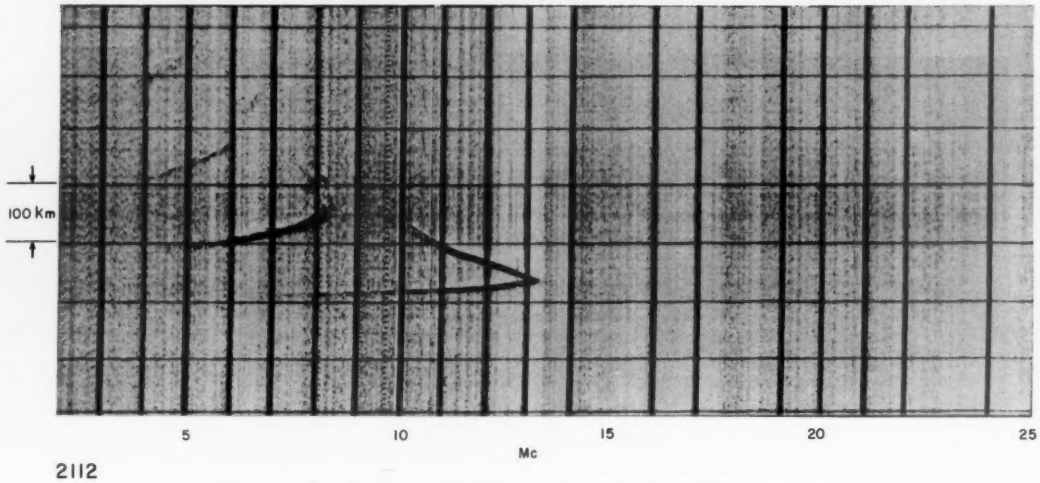


FIGURE 16. *Sterling to Boulder sweeps with MUF extensions.*



2112

FIGURE 17. *Sterling to Boulder sweep without the MUF extension.*

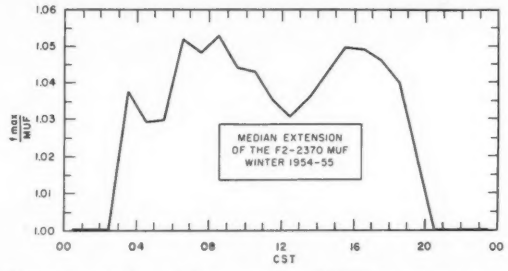
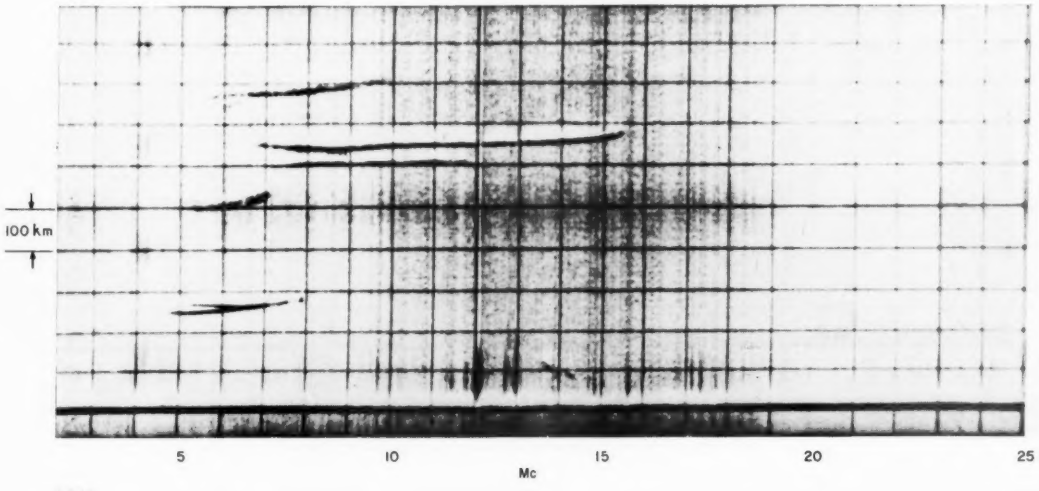


FIGURE 18. *The winter diurnal variation of the percentage MUF extension for the Sterling to Boulder path.*

NOVEMBER 18, 1952



1411

FIGURE 19. *Winter sweep showing the absence of the Pedersen ray signal over the Sterling to St. Louis path.*

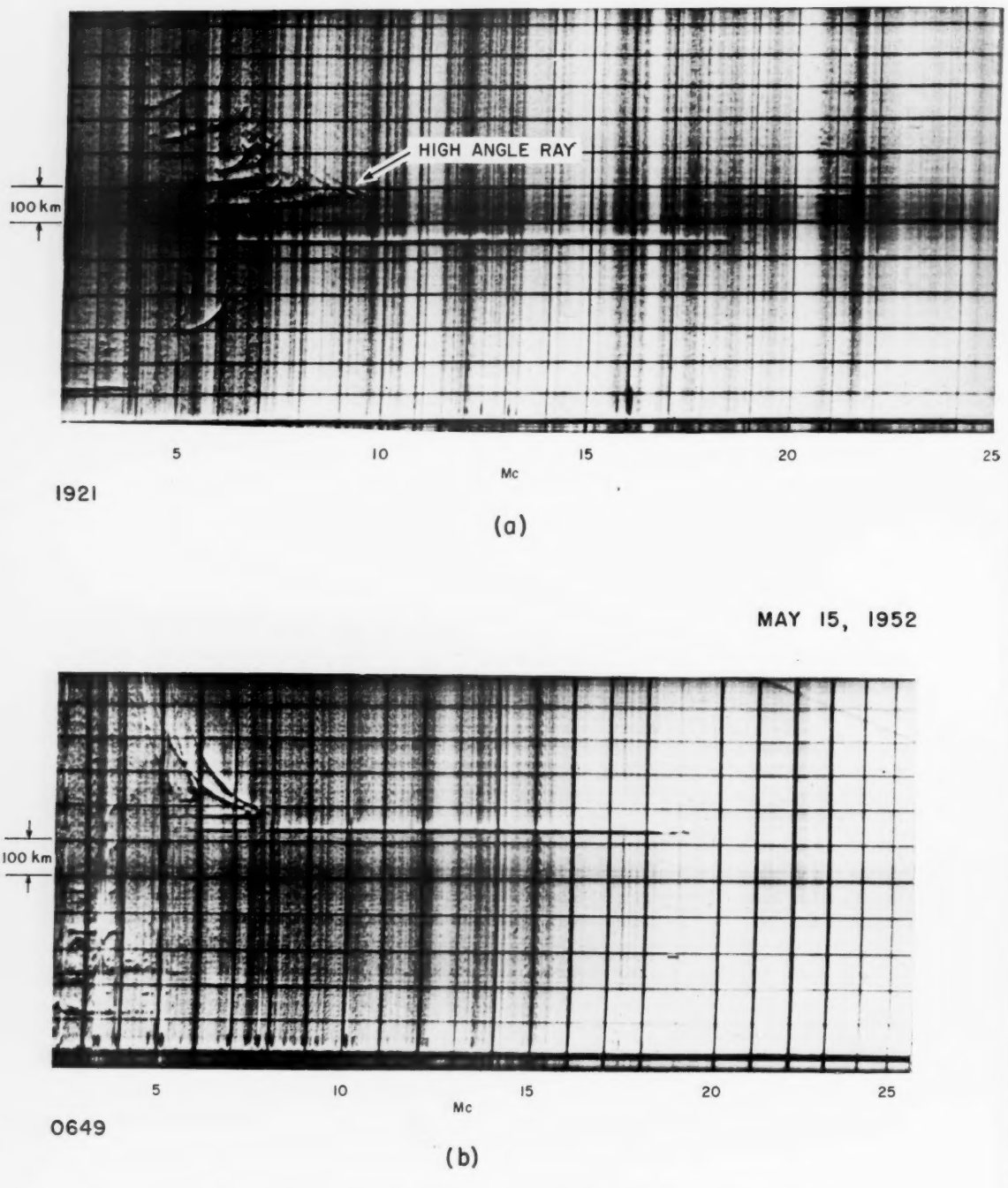


FIGURE 20. *Sterling to St. Louis sweeps with examples of relatively strong Pedersen rays.*

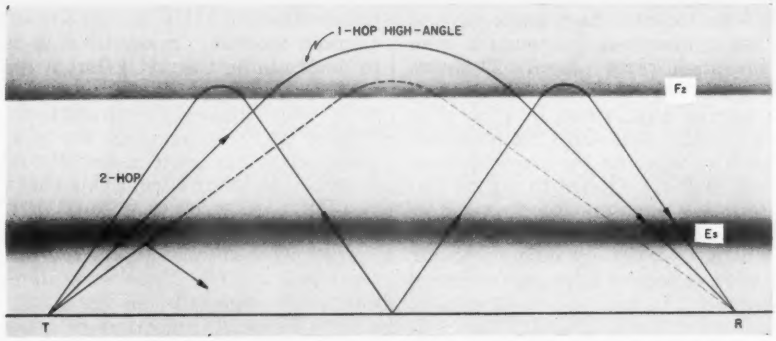
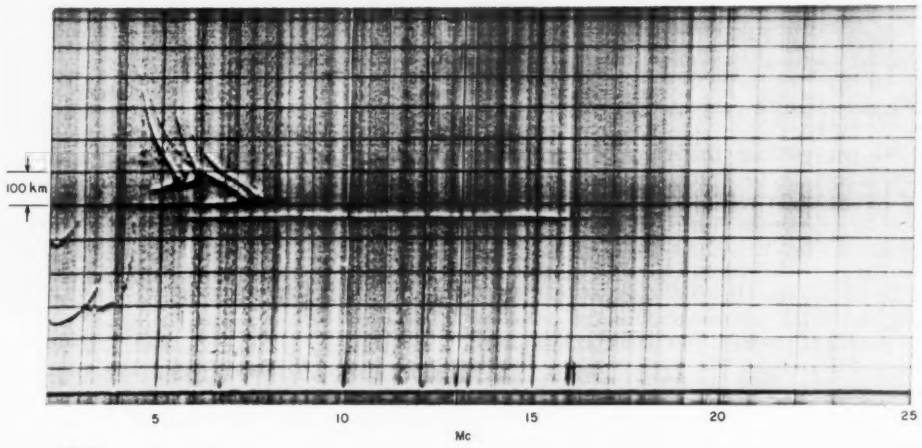


FIGURE 21. *Illustration of the way in which the 1-hop low-angle ray may be cut off by the E layer while the Pedersen ray and the 2-hop mode can propagate between transmitter (T) and receiver (R).*

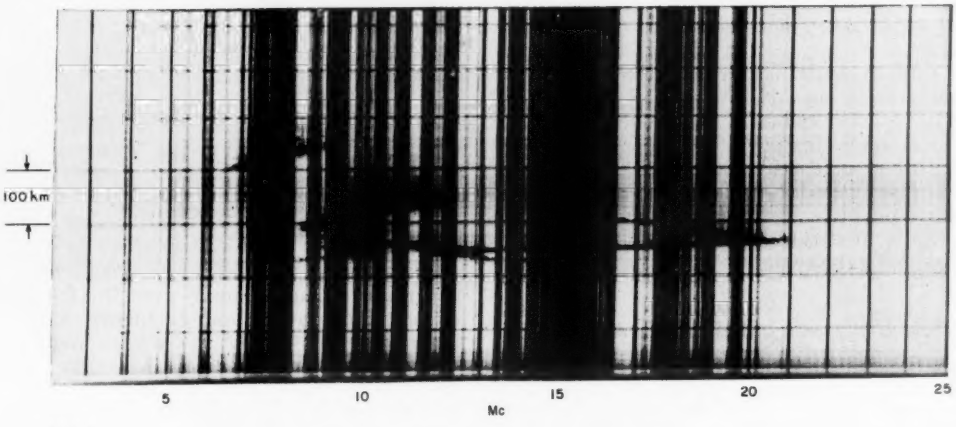
MAY 15, 1952



0637

FIGURE 22. *Low-frequency cut-off on the Sterling to St. Louis path.*

DECEMBER 20, 1954



1024

FIGURE 23. *Low-frequency cut-off on the Sterling to Boulder path.*

4.7 Mc. It will be seen that the high-angle rays of both modes have lower observed frequencies than the corresponding low-angle rays. Figure 23 shows the same phenomenon for the 2,370-km path. We see that the trace having the lowest frequency at which echoes appear is the three-hop $F2$ at a frequency of about 6.8 Mc, whereas the one-hop echo first appears at about 9 Mc. (Refer to figure 21.) Sulzer [15] attributed this phenomenon to antenna patterns but the lowest observed frequencies are in good agreement with the E-layer cut-off hypothesis, and we conclude that the effect of antenna pattern is of secondary importance.

c. Separation of the Ordinary and Extraordinary Rays at the MUF

As the distance of transmission is increased, the maximum frequencies of both the ordinary and the extraordinary waves are increased. However, the separation between the two MUF's decreases with increase of distance. This frequency separation depends upon the angle of incidence; the frequency; the strength of the magnetic field; and, to some extent, on the direction of propagation with respect to the magnetic field.

In figure 24a we see that the effect of refraction in the lower layers is to increase the frequency separation of the ordinary and extraordinary MUF's. This is a result of a decrease in the angle of incidence of the ray at the $F2$ layer. The separation is about 0.6 Mc.

In figure 24b we see the effect of angle of incidence by comparing the MUF separations for the one-hop and two-hop modes of propagation where the separations are about 0.3 and 0.6 Mc, respectively. Part of this change may be due to the lower maximum frequencies of the two-hop signals since even at vertical incidence the critical frequency separations between the ordinary and extraordinary rays increase as the critical frequencies decrease. This argument cannot be applied to the case in figure 24a since the $F2$-layer MUF's in both figures 24a and 24b are practically equal (around 7.2 Mc). Note in the one-hop mode how for a given ground range the frequency separation increases with increase of group path.

A statistical analysis has been made to find the mean separation of the MUF's for east-west propagation in the United States. The separations refer to the one-hop mode for the 2,370-km path, the one-hop and two-hop modes for the 1,150-km path and the vertical-incidence data. The results are plotted in figure 25, where it is seen that, on the average, the separation falls off steadily with distance from about 0.8 Mc at vertical incidence to about 0.2 Mc at 2,500 km.

3.4. Discussion of the NBS Results

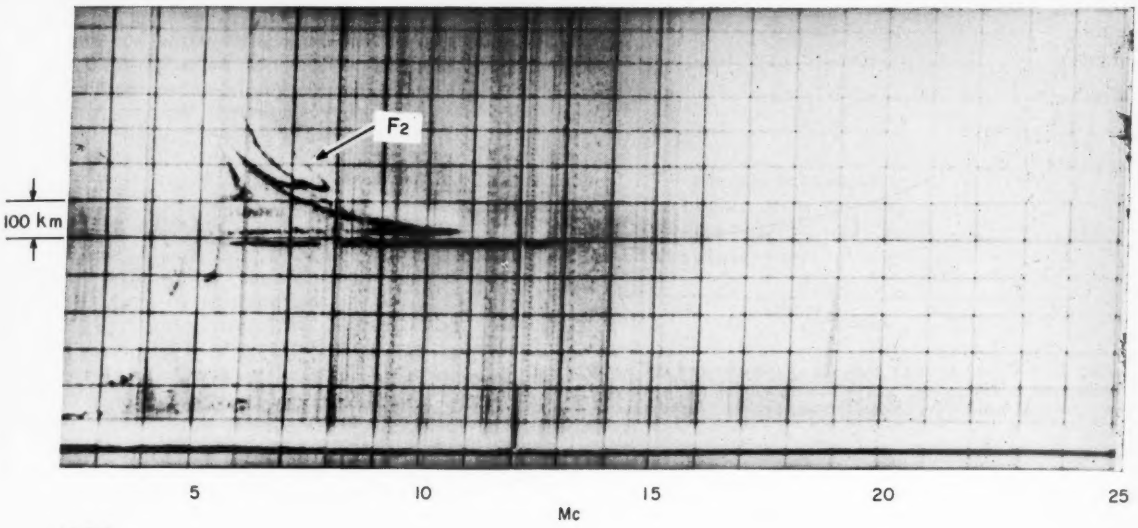
The experiments described above have shown that over the distances of transmission concerned (1,150 and 2,370 km) the $F2$ layer is not always the layer which determines the MUF. Sporadic-E is of particular importance on the shorter path and usually determines the MUF during the sunlit periods of the summer months. Sporadic-E is not only important in determining the MUF but it may also decide the lowest observed frequency because of cut-off. Thus it is clear that MUF predictions based solely on regular layer propagation are of limited value over paths of the order of a few thousand kilometers in length. It follows from this that tests of the accuracy of regular layer prediction methods based on times of fade-in and fade-out of CW signals may be highly misleading. Again, during the summer daylight periods, the $F1$ layer is important in deciding the MUF especially on the 2,370-km path because the one-hop E reflection is generally nonexistent. The $F1$ layer is relatively more important during periods of low sunspot activity when daytime $foF2$ sinks to relatively low values and also during magnetically disturbed periods.

The $F2$ layer is, however, the dominating layer by night at all seasons and throughout the winter day except for short periods when sporadic-E may be present.

Perhaps the most important aspect of the oblique-incidence experiments is the relationship between vertical-incidence data and oblique-incidence data. The first feature of interest is the close correlation between the temporal changes (both long and short term) between the observed MUF's and the calculated MUF's. The calculated MUF's are, on the average, some 3 percent lower than the observed values. This percentage difference is fairly consistent especially during the winter months. The source of this systematic error must lie in the assumptions in the theory underlying the Smith transmission curves. The essential assumptions made in the calculation of these curves are: 1. The geometric optical approximation is valid. 2. The influence of the earth's magnetic field can be neglected. 3. The effect of ionosphere curvature can be accounted for by a correction factor k which varies with distance of transmission only. 4. The ionosphere is spherically stratified. 5. Tropospheric refraction is negligible. The first assumption, i.e., the applicability of geometric optics, probably needs some modification near the MUF. This is especially so in the case of pulse propagation where the signal is composed of components spread over a fairly wide frequency range. For instance, when the carrier frequency is at the ray MUF, half the frequency components have frequencies above the MUF. Some modification of the variation of amplitude with frequency near the MUF will result from the application of wave theory but the effect on the value of the MUF is probably slight.

The influence of the earth's magnetic field on the ray path is usually small over the frequency range covered by the oblique-incidence work described above, except during the winter nights when the MUF falls to around 3 to 4 Mc. At these frequencies the frequency separation between the ordinary and extraordinary MUF's shows a noticeable increase over the separation at the higher frequencies. The important thing about the effect of the magnetic

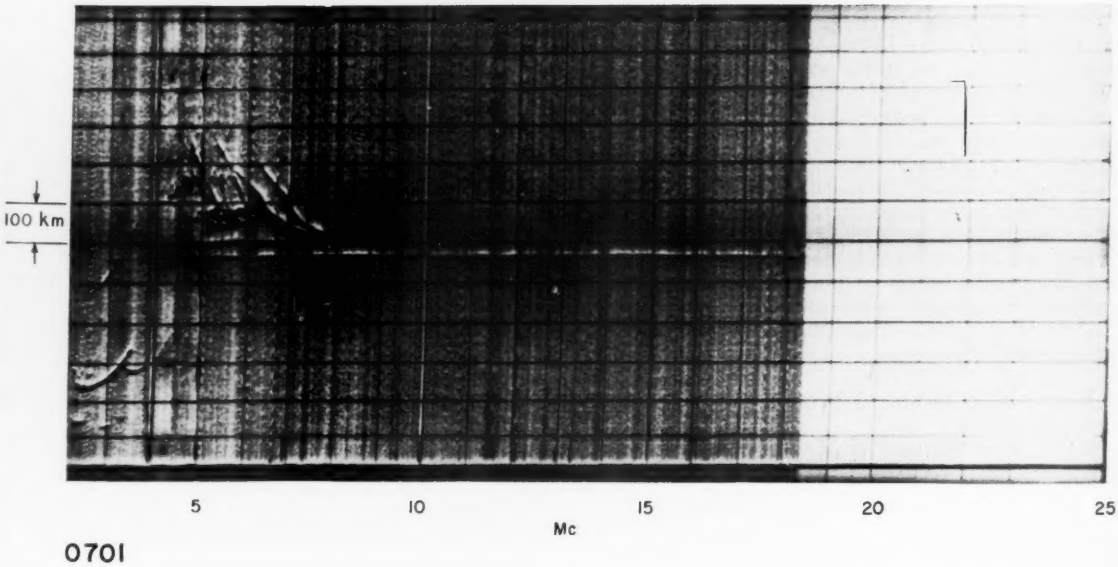
SEPTEMBER 26, 1952



1120

(a)

MAY 15, 1952



0701

(b)

FIGURE 24. *Sweeps showing the dependence of the frequency difference between the ordinary and extraordinary MUF's on group retardation and angle of incidence for the Sterling to St. Louis path.*

Note how the frequency separation increases as the virtual height increases.

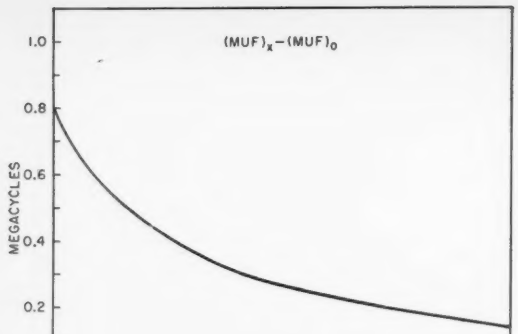


FIGURE 25. *Variation with distance of the observed difference between the ordinary and extraordinary ray MUF's for east-west propagation in the United States.*

field, however, is that it always tends to increase the ground range of the ordinary wave as compared with the case with no magnetic field. In other words, the effect of the field is always to reduce the MUF which is contrary to the observed phenomenon. The only exception to this is the case of east-west propagation along the magnetic equator where the MUF is unaffected by the magnetic field.

Wieder [14] has suggested magnetoionic deviation coupled with short-term variations in ion densities (and their geographical and temporal distributions) as a possible cause of the dispersion of the differences existing between the calculated and measured MUF's. Such an explanation holds, of course, only if the point of reflection for the oblique ray is located at an appreciable distance from that for the vertically incident ray (at an equivalent frequency) and if, in addition, appreciable horizontal gradients in ion density sometimes exist in the distance that separates the reflection points.

It is difficult, because of the number of parameters involved, to state with any generality the magnitude of the effects of magnetoionic deviation. However, insight into the general problem is given by application of Booker's [24] semigraphical method for the following conditions:

a. Horizontally stratified parabolic layer of 100-km semithickness;
b. Gyrofrequency, f_H=1.5 Mc;
c. Magnetic dip=70°;
d. Vertical plane containing direction of phase propagation perpendicular to vertical plane containing magnetic field direction (east-west propagation);
e. Angle of incidence for oblique ray, θ=60°;
f. Frequency (f), propagated obliquely=15 Mc (f_H/f=0.1); and
g. Vertical incidence (equivalent) frequency, $f \cos\theta$=7.5 Mc (f_H/f=0.2).

The results of the computation of the magnetoionic deviation for layers having ordinary wave critical frequencies (f_0) of 7.5 and 10 Mc are given in table 1.

TABLE 1

	Ordinary ray				Extraordinary ray			
	Vertical incidence		Oblique incidence		Vertical incidence		Oblique incidence	
	H	*D*	*H*	*D*	*H*	*D*	*H*	*D*
f_0=7.5.....Mc..	100	46	100	2.5	55	2.8	88.5	1.5
f_0=10.....Mc..	33.9	4.3	33.9	0.5	25.8	1.2	31.8	0.4

H—height of reflection in kilometers above the bottom of the layer.
D—distance in kilometers of reflection point from the vertical plane containing the ray incident on the layer.

The first case listed (f_0=7.5 Mc) is one which cannot arise in practice; i.e., reflection of the oblique ray from the level of maximum ion density. For the flat earth, this corresponds to an infinitely great transmission distance. But it is notable that even here, the deviation of the obliquely incident ray is probably negligible (2.5 km computed), although the deviation of the corresponding vertically incident ray is not. The second case is of greater interest from a practical standpoint (reflection well below the level of maximum ionization), and here it is seen that neither ray (vertical or oblique) is deviated far from the plane of incidence. Although, in a distance of about 4 km (difference in computed deviations for the vertical and oblique rays at reflection), disturbing turbulence or gradients may well exist, these same disturbing elements could seriously affect the ray paths even if the effects of the magnetic field were neglected. The fact that the ionosphere is not uniform, i.e., not horizontally, or spherically stratified (the oblique ray "samples" this nonuniform medium over a relatively great distance), appears to be more pertinent to an explanation of the disagreement (or the variations in the amount of disagreement) between calculated and observed MUF than the magnetoionic deviation.

Interest in the magnetoionic deviation is then, more-or-less, academic, but it may be worthwhile to summarize the results of computations based on Booker's method. For east-west propagation, as the penetration into a given layer increases, either because of decreasing angle of incidence or increasing frequency, the amount of magnetoionic deviation increases. The extraordinary ray is deviated less than the ordinary at all latitudes and for all angles of incidence (except at the equator where neither ray is deviated out of the plane of incidence). For north-south propagation (in the vertical plane containing the magnetic field direction) no magnetoionic deviation out of this plane is experienced, but the path in the ionosphere is asymmetric, reflection points being toward the pole (ordinary) or toward the equator (extraordinary) from the path midpoint by distances of the same order as the lateral deviations suffered over an east-west path. Also, for north-south propagation the ordinary ray is reflected

from a higher level and the extraordinary from a lower level than is the case for an east-west path. This effect, although probably relatively unimportant, should be kept in mind when comparisons are made between north-south and east-west paths.

The possible error in the transmission curve k factor was discussed by Wieder, who concluded that the average discrepancy of 3 percent was unlikely to be due to the k error. This conclusion is, however, somewhat uncertain and it is of interest to consider this point a little further. The k factor depends in a complicated way on the electron distribution in the ionosphere and varies, to some extent, with the height of reflection and angle of incidence. The k is, therefore, not constant for a given distance of transmission as was assumed in the theory underlying the transmission curves. The variation of the ratio (f/f') with virtual height of reflection is shown in figure 26, where f' is the vertical-incidence frequency equivalent to a frequency f at oblique incidence. This figure, due to Wieder, shows that for low virtual heights of reflection the empirical values of f/f' are larger than sec Φ_0, while at high virtual heights this effect tends to diminish. It will be noticed that the high values of virtual heights were obtained during the summer and are brought about by group retardation in the lower layers. They do not, therefore, represent the true heights of reflection.

One of the assumptions on which the transmission curves are based is that there are no horizontal gradients of electron density in the ionosphere. There is now considerable evidence for both large scale and small scale irregularities in the ionosphere, the large-scale gradients being known as tilts. Wieder, in discussing the effect of a north-south gradient on the calculation of MUF's, pointed out that "the northward deviated vertical-incidence ordinary wave encounters systematically lower densities than the obliquely incidence wave." We have seen above, however, that the lateral deviation on the equivalent vertical-incidence frequency is only about 4 km greater than the lateral deviation of the oblique ray at an angle of incidence of 60°. Hence, it appears unlikely that this would explain the observed discrepancy.

It has also been suggested that tilts might account for the discrepancy, but if the presence of a tilt shifts the point of reflection of the obliquely incident ray, then it should also produce a corresponding shift in the point of reflection of the vertically incident ray on the equivalent frequency.

Another point of considerable theoretical importance is the origin of the so-called "MUF extension". D. K. Bailey has suggested that the extension is the result of a modification in the angle of incidence of the ray on the F layer due to scattering of energy by irregularities in the E region. Such a mechanism would account for signals on frequencies above the classical MUF as would reflection from a rough F layer. One might expect some correlation with sporadic E on this hypothesis; so far very little correlation has been found.

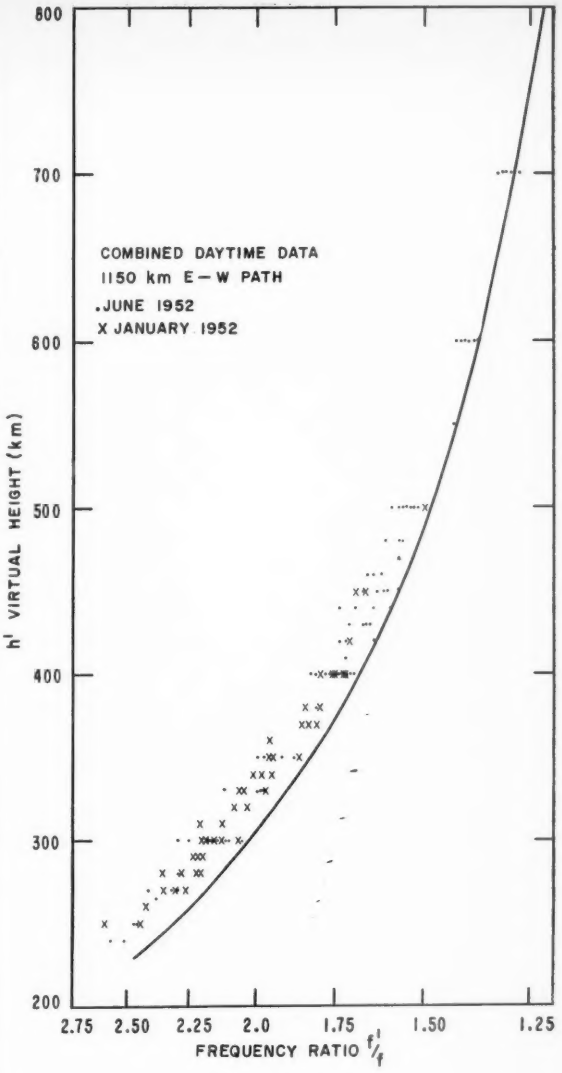


FIGURE 26. *Graph illustrating the variation of the observed ratio of f'/f with equivalent height over the Sterling to St. Louis path.*

$f' \equiv$ the vertical-incidence frequency equivalent to the frequency f. The full line is the sec Φ_0 line.

4. Conclusions

The results of the oblique-incidence experiments at the Bureau are in general agreement with those made by other laboratories. In particular, the discrepancy of about 3 percent between the observed and calculated MUF is similar to that found elsewhere, except that no marked percentage increase appears to occur with increase of distance as reported by the Japanese [16]. The latter probably refer to

the maximum observed frequency, whereas the former refer to the classical MUF. The relative importance of the high-angle ray during summer has been confirmed in nearly all the other published work on the subject.

5. References

[1] F. T. Farmer and J. A. Ratcliffe, Wireless waves reflected from the ionosphere at oblique incidence, Proc. Phys. Soc. (London) **48,** 839 (1936).

[2] F. T. Farmer, C. B. Childs, and A. Cowie, Critical frequency measurements of wireless waves reflected obliquely from the ionosphere, Proc. Phys. Soc. (London) **50,** 767 (1938).

[3] T. L. Eckersley, S. Falloon, F. T. Farmer, and W. O. Agar, Wireless propagation and the reciprocity law, Nature **145,** 222 (1940).

[4] W. J. G. Beynon, Propagation of radio waves, Wireless Eng. **25,** 322 (1948).

[5] W. Dieminger, K. H. Giesweid, and H. G. Möller, Echolotungen der ionosphäre mit veränderlicher frequenz bei schragem Einfall, NTZ-Nachr tech. Z. **8,** 578 (1955).

[6] W. Dieminger and H. G. Möller, Echo sounding experiments with variable frequency at oblique incidence, Nuovo cimento [10] **4;** Suppl. No. 4 (1956).

[7] H. G. Möller, Sweep-frequency oblique-incidence experiments over a distance of 1320 km, J. atmospheric and Terrest. Phys. **9,** 155 (1956).

[8] H. G. Möller, Further results of sweep frequency oblique-incidence experiments, J. Atmospheric and Terrest. Phys. **13,** 173 (1958).

[9] W. Dieminger, H. G. Möller, and G. Rose, Long distance single *F*-hop transmissions, J. Atmospheric and Terrest. Phys. **13,** 191 (1958).

[10] J. W. Cox and Kenneth Davies, Oblique incidence pulse transmission, Wireless Eng. **32,** 35 (1955).

[11] J. H. Chapman, Kenneth Davies, and C. A. Littlewood, Radio observations of the ionosphere at oblique incidence, Can. J. Phys. **33,** 713 (1955).

[12] E. Warren and E. L. Hagg, Single-hop propagation of radio waves to a distance of 5300 km, Nature **181,** 34 (1958).

[13] P. G. Sulzer and E. E. Ferguson, Sweep-frequency oblique-incidence ionospheric measurements over an 1150-km path, Proc. I. R. E. **40,** 1124 (1952).

[14] B. Wieder, Some results of a sweep-frequency propagation experiment over an 1150-km east-west path, J. Geophys. Research **60,** 395 (1955).

[15] P. G. Sulzer, Sweep-frequency pulse-transmission measurements over a 2,400-km path, J. Geophys. Research **60,** 411 (1955).

[16] Unpublished, see C.C.I.R. Doc VI/53E (June 1958).

[17] F. Kift, Single-hop propagation of radio waves to a distance of 5,300 km, Nature **181,** 1459 (1958).

[18] Ionospheric radio propagation, ch 6, NBS Circ. 462 (1948).

[19] E. V. Appleton and W. J. G. Beynon, The application of ionospheric data to radio communication problems, Part I, Proc. Phys. Soc. (London) **52,** 518 (1940).

[20] N. Smith, The relation of radio sky-wave transmission to ionospheric measurements, Proc. IRE **27,** 332 (1939).

[21] J. W. Wright, R. W. Knecht, and K. Davies, Ann. IGY **III,** Pt. I (1957).

[22] R. Silberstein, Sweep-frequency backscatter—some observations and deductions, IRE Trans. **AP–2,** 56 (1954).

[23] R. Silberstein, The use of sweep-frequency backscatter data for determining oblique-incidence ionospheric characteristics, J. Geophys. Research **63,** 335 (1958).

[24] H. G. Booker, Application of the magneto-ionic theory to radio waves incident obliquely upon a horizontally stratified ionosphere, J. Geophys. Research **54,** 243 (1949).

BOULDER, COLO.

(Paper 63D2–15)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Fields in Electrically Short Ground Systems: An Experimental Study

A. N. Smith[1] and T. E. Devaney[1]

(April 20, 1959)

An experimental study of magnetic field distribution in a simplified radial ground system on poorly conducting soil under an electrically short, toploaded monopole is described. It is shown that the distribution is that expected from the theory of J. R. Wait in those portions of the radial system satisfying the assumptions of the theory, and that the theory may still be successfully applied for H-field power loss computations even when this is not fully the case. The particular model system studied exhibits a condition suggesting damped standing waves on the radials in the area where the radial spacing exceeds that required by the theory.

1. Introduction

For an electrically small monopole, the ground-system power losses very often are a controlling factor in radiation efficiency. The decision to construct at Cutler, Maine, a new high-power, very-low-frequency station with specified efficiency greatly exceeding that of any such existing station has thus resulted in a large amount of recent interest in ground-system design.

The final ground-system configuration for Cutler came basically from computations based on Wait's [1, 2, 3] equivalent surface impedance formula for radial current (H-field) power loss for a plane, radial grid of wires laid with perfect contact on an earth-air interface; and from design modifications brought about according to ground charging current (E-field) power loss computations based on formulas due to Abbott [4], Wheeler [5], and Wait [6]. The soundness of the theoretical design was partially and indirectly confirmed at the U.S. Navy Electronics Laboratory, San Diego, Calif., by comparing losses derived from impedance measurements [7] with those computed from theory for a model toploaded vlf monopole over a radial-wire system and also over a perfectly conducting ground [8]. A somewhat more direct test of theory was made by W. G. Hutton and others of Smith Electronics, Inc., Cleveland, Ohio [9, 10], who measured the field distribution (magnitude only) in a full-scale, fan-shaped ground-system sector constructed at the Cutler site. An analysis of his data by Wait [11] showed that the measured field distribution did indeed generally agree with that expected from a division of ground system current according to the surface impedance ratio derived from the theory.

However, there existed sufficient uncertainty about some conditions in all these experiments that a refined field-distribution study seemed desirable. Accordingly, a phase- and magnitude-measurement program was carried out on the vlf model previously used for impedance measurements. The dual objective was to test the validity of the theory and to investigate more adequately the direction and degree of departures therefrom those existing in the model.

[1] U.S. Navy Electronics Laboratory, San Diego, Calif.

2. Theoretical Discussion

The working formula for ground loss R_g in terms of the azimuthal magnetic field $H_\phi(\rho,\theta)$ at the surface of the ground ($\theta=0$) referred to the antenna feed point where current I_o flows is

$$R_g=\text{Real part of } \frac{1}{I_0^2}\int_{r_i}^{\infty} Z[H_\phi(\rho,0)]^2 2\pi\rho d\rho. \quad (1)$$

This equation is derived [2, 3] from the incremental Poynting vector flux due to the tangential component (E_ρ) of electric field arising from the finite conductivity of the soil. In the derivation use is made of an approximate boundary condition through which the flux is expressed in terms finally of H_ϕ, the magnetic field of the monopole over an infinitely conducting ground. For convenience the integration is performed over two regions, $r_i \leq \rho \leq r$, and $\rho \geq r$, r being the length of the wire radials, since, in the area occupied by the ground system, the surface impedance Z is the combination of wire impedance Z_w and soil impedance Z_s and in the remainder it is simply Z_s. r_i is the inner extremity of the radials.

The five major assumptions made in deriving equation (1) are stated below.

(1) The approximate boundary condition $E_\rho = ZH_\phi$ is used, requiring that ρH change slowly in a distance $d=|\gamma^{-1}|$ where $\gamma=[i\mu\omega(\sigma+i\omega k\epsilon_o)]^{1/2}$ is the propagation constant of the soil.

(2) The combined surface impedance varies slowly in $|\gamma^{-1}|$.

(3) The radial wire screen of N radials in any small region a distance ρ from the antenna base equals that of a grid of parallel wires of the same mean spacing $s=2\pi\rho/N$, i.e., $Z_w=(i\eta_o s/\lambda)\ \ln\ (s/2\eta c)$ where $\eta_o=377$ ohms, and c is the wire cross section radius. This requires the condition $s \ll |\gamma_e^{-1}|$ where

$$\gamma_e=\left(\frac{\gamma_0^2+\gamma^2}{2}\right)^{1/2}$$ is the effective propagation constant

of the composite medium at the interface ($\gamma_o=i\beta$).

(4) To replace the actual magnetic field H_ϕ by H_ϕ^∞ at the surface of the ground it is necessary that

$|\gamma^{-1}| \ll 1/\beta$, i.e., the soil skin depth be much less than a free-space electrical radian.

(5) The surface impedances of screen and soil must combine as simple parallel impedances, i.e., mutual coupling and potential differences between wires and soil are zero, so that

$$Z = Z_s Z_w/(Z_s + Z_w) \tag{2}$$

within the ground system.

These assumptions must be satisfied in setting up any direct experimental test of equation (1).

Since this paper is largely concerned with the field structure due to division of current in the ground system according to assumption (5), its implications are now developed. The total ring current crossing a cylindrical boundary of radius ρ drawn about the monopole base as a center is $I_t = I_w + I_s = 2\pi\rho H_\phi$. Normalizing with respect to base current ($h_\phi = H_\phi/I_0$) and removing the factor $2\pi\rho$, assumption (5) leads to

$$h_\phi = h_\phi Z_s/(Z_w + Z_s) + h_\phi Z_w/(Z_w + Z_s). \tag{3}$$

The first term on the right of (3) corresponds to wire current I_w. Letting $I_w/I_0 = Ni_w$, so that i_w is the normalized current in one radial wire, the I_w term is seen to be exactly $Ni_w/2\pi\rho$. Hence the magnetic field h_w at a field point d meters from the wire center is

$$h_w = h_\phi \rho Z_s/Nd(Z_w + Z_s). \tag{4}$$

The second term on the right of (3) corresponds to current in the soil I_s, and in fact is $i_s/2\pi\rho$, where $i_s = I_s/I_0$. If the wire centers lie exactly in the plane of the earth-air interface, then the magnetic field tangent to it, h_s, can be simply evaluated by applying Ampere's law to the path illustrated in figure 1, where n, the number of skin depths δ from the surface to the portion def of the path, is so large that H_ϕ along it is negligibly small. At the interface there is no tangential contribution from the wires themselves, and along the vertical portion the contributions cancel. Hence

$$2\pi\rho h_s = Ni_w/2 + i_s. \tag{5}$$

Subtracting this from (3), it is seen that

$$h_s = \frac{h_\phi}{2}\left(1 + \frac{Z_w}{Z_s + Z_w}\right). \tag{6}$$

That is, in the plane of the wire centers the total tangential field is the incident field of the downlead plus a contribution due to the induced current flowing in the soil. The wires carry only a portion of the image current. In the special case of a wire counterpoise and no soil, the second term vanishes, and all the image current flows in the wires. In the case of infinitely conducting ground, on the other hand, the second term also equals $h_\phi/2$ (perfect reflection), the distinction between h_w and h_s is lost, and the total field is just h_ϕ.

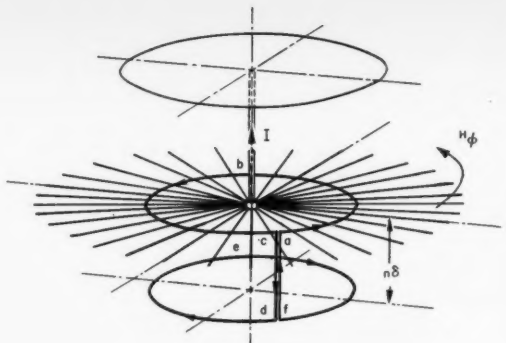


FIGURE 1. *Integration path for evaluating* h_s.

From the above considerations one sees that to determine the validity of the theory of H-field power losses, it is sufficient to observe that the fields h_w and h_s satisfy equations (4) and (6) when assumptions (1), (2), (3) and (4) are met.

3. Experimental Procedure

In the Cutler design [11] and in the NEL models used for the impedance measurements, soil constants and ground system parameters were such that assumptions (1), (2), and (3) were certainly met except at points either very close to the downlead or within a soil skin depth of the radial ends. For the purpose of the present experiment, condition (4) was regarded as an object of interest, while condition (5) was the main point at issue.

Preliminary computations indicated that there would be sufficient individual wire current to yield usable signal for the instrumentation employed only if relatively few radials made up the ground system. This condition was required also to insure that the signal from a field point over a wire would be large compared to that from the nearby soil current, and moreover to insure that the wire field would not contain significant contribution from adjacent wires. To a radius of 100 ft the soil surface to a depth of several inches was stripped of all old radials, vegetation, and inhomogeneities, and was regraded. Eight No. 30 bare copper wires 40 ft long were laid, stretched, and covered with a thin layer of mud to make contact with the interface as ideal as possible. The inner termination was a 1-ft-radius brass plate

grounded to six 18-in. copperweld rods; the outer terminations were ¼ in. by 6 in. galvanized boat nails. Later, another set of eight radials was run in between the originals. Finally all were extended to 80 ft. The antenna was otherwise left as it had been; it is a somewhat conical, 36-radial toploaded supported at its periphery by a hexagonal spreader stretched over six 8-ft insulated galvanized iron pipe towers symmetrically spaced on a 13-ft-radius circle; the central support is the 12½-ft downlead. At 1.5 Mc, the only frequency employed for both phase and magnitude measurements (an extensive series of magnitude distribution studies was made at 3.0 as well as 1.5), this structure on an ideally conducting ground has a base impedance of $(0.40-j290)$ ohms, and has an effective height of 7½ ft.

Figure 2 shows the measurement scheme. It comprised two 1-in.-radius, shielded, 20-turn pickup coils C_1 and C_2 connected by their coax lines and trimmers T_1 and T_2 through similar d-c amplifiers A_1 and A_2 to the deflection system of an oscilloscope. Coil C_1 was kept at the reference point 1 ft from the downlead over a radial wire, while C_2 was the exploring coil. Both coils were oriented for maximum coupling, and the shields of both were insulated from direct contact with soil or wires. An RC phase shifter could be introduced into the circuit of one of the coils in order to sense the absolute direction of an observed phase shift. Base current was monitored by a calibrated milliammeter. The sensitivity of the field-detection system was obtained by comparison with readings on a field-strength meter (f-s meter No. 1) set up beyond the radials, and also by measurement of the induction field of a known, long, straight current filament. Harmonic content of the antenna current was checked as a part of the calibration procedure.

Since the distance d from wire to field point was of the same order as the loop diameter ($d=1.94$ in.) the field cannot be considered uniform over the coil area. The average field h_w actually measured is related to the value at the field point by a factor

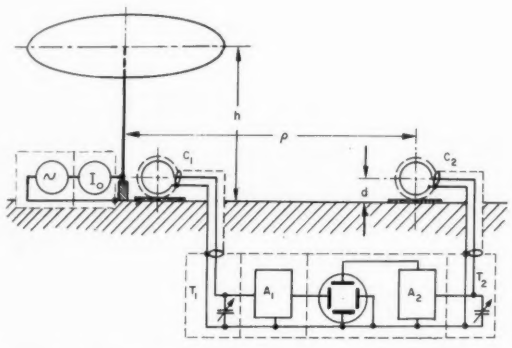


FIGURE 2. *Scheme for measuring field magnitudes and phases.*

$h_w=Fh_w$ where for coplanar radial and coil turns

$$F=\frac{b^2\sqrt{1-b^2/d^2}}{2d^2(1-\sqrt{1-b^2/d^2})}. \tag{7}$$

For the coils used this amounted to 0.80. Calibration against f-s meter No. 1 gives $h_w=KS_w/I_0$, where S_w is the deflection developed on the oscilloscope, and K is the calibration figure. Stability of K turned out to be to within a percent over an extended period.

For measuring the field between the wires, small trenches were dug and the coil centers alined in the plane of the wires within one-fourth in. These measurements were done largely along the bisectors of the angle between adjacent wires, but exploration was also carried out azimuthally for several fixed values of ρ to test for constancy of field between wires.

The test of equivalence of H_ϕ and H_ϕ^∞ was a comparison between the near-zone field measured for the antenna over the brass center plate only and the values computed from Wait's formula [12]

$$H_\phi^\infty=\frac{I_0e^{-i\beta\rho}}{2\pi\rho}\left[\left(\frac{h^2}{h^2+\rho^2}\right)^{1/2}+i\beta\rho\tan^{-1}(h/\rho)\right]+D \tag{8}$$

where D is the contribution to H from the toploaded disk, and h is the height of the disk above the ground (also taken to be the effective height). This was done using an untuned probe with a second field-strength meter calibrated against f-s meter No. 1. Later a partial confirmation was had from a toploaded monopole set up over a small ground system half of which was hardware-cloth mesh, the other half consisting of 25 radials.

Finally, it was necessary to know something about the electrical characteristics of the soil. To determine these, the properties of short, known lengths of buried open-wire transmission lines of several spacings were measured at the frequencies of interest. From this information the electrical constants of the medium were derived; the best values for 1.5 Mc appeared to be conductivity $\sigma=0.010$ mho/m, dielectric constant $k=50$.

4. Discussion of Results

Figure 3 shows the results for the test of assumption (4). In general, the agreement between the measured h_ϕ and theoretical h_ϕ^∞ is good. The departure within the periphery of the toploaded is not unexpected inasmuch as (a) the toploaded was somewhat conical rather than flat as assumed in computing h_ϕ^∞, and (b) the soil constants imply a skin depth of about 12 ft at 1.5 Mc, so that at no point under the toploaded can the currents be regarded as flowing in a plane, radial sheet without vertical divergence, which has the effect of decreasing the apparent field. The increase in h_ϕ over the theoretical curve beyond the toploaded seems to be mostly due to an 8 percent calibration error discovered in

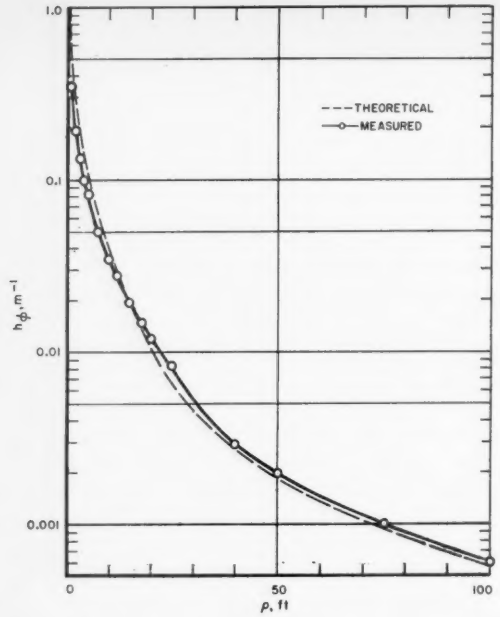


FIGURE 3. *Comparison of measured and theoretical* hϕ *for homogenous, poorly conducting ground.*

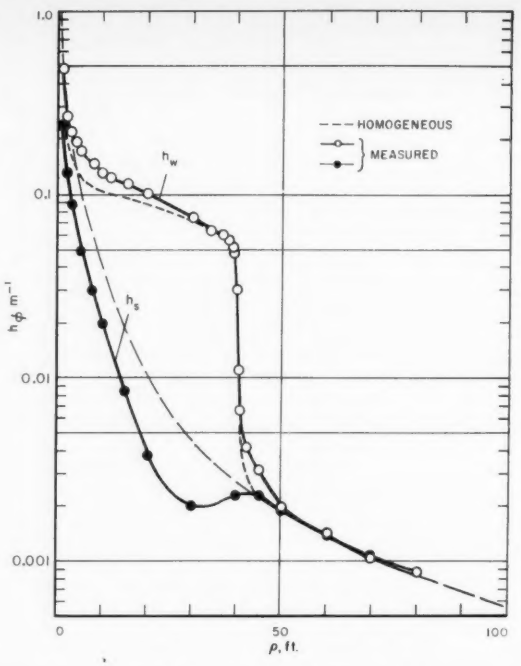


FIGURE 4. *Measured field magnitudes over and between wires of 40-ft 16-radial system.*

Short-dashed curve is $|h_w|-|h_s|$, representing minimum possible h_w.

f-s meter No. 1 after the conclusion of the experimental program. Thus for soils no poorer, relatively speaking, than that in this experiment, confirmation seems to be had of the propriety of replacing h_ϕ by h_ϕ^∞ in H-field power loss computations for the region beyond the ground system.

Analysis of the radial ground system experiments proceeds from the observation that there seems to be only qualitative agreement between theory and experiment, and a significant difference between phenomena in the 40-ft system and those in the 80-ft (see figs. 4, 5, and 6). It is first in order to determine the degree to which these somewhat sketchy ground systems satisfy the assumptions of the theory.

Since ρH changes at most by a factor of two in the ring $\delta \leq \rho \leq 2\delta$ (δ being the skin depth) and less rapidly thereafter, the first assumption appears to be met fairly well except for the area under the top-load. The second assumption also appears to be satisfied for the same range of ρ for the 16-radial ground system, but not for the eight. The third assumption was not satisfied by the 80-ft 16-radial system for $\rho >$ 40 ft; only poorly in 20 ft $< \rho <$ 40 ft, for both 40- and 80-ft systems; fairly well in $\rho <$ 20 ft. In the eight-radial system it was not satisfied. Assumption (4) seemed to be well satisfied not only from the fact that at 1.5 Mc $|\gamma| = 10\beta$, but also from the experimental confirmation discussed above. Finally, assumption (5) seemed to be only approximately satisfied. Azimuthal sweeps at constant ρ over a quadrant of the ground system showed constancy in $|h_s|$ within a few percent for all locations except those within a horizontal distance of 0.1 δ to

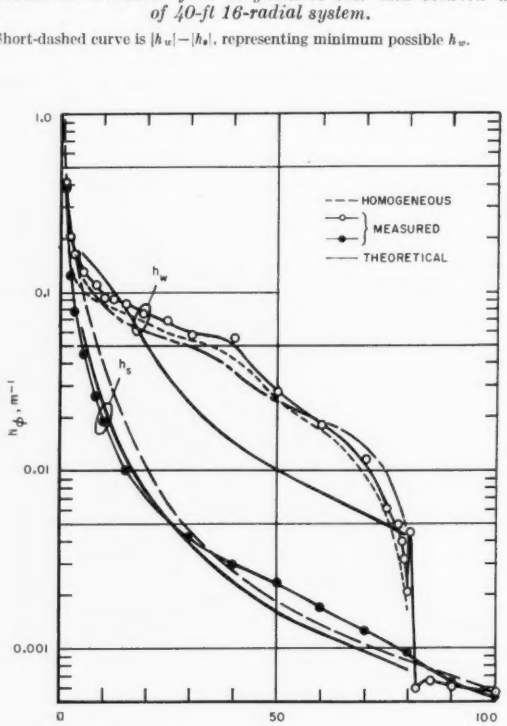


FIGURE 5. *Comparison of measured and theoretical magnitudes over and between wires of 80-ft 16-radial system.*

Dash-dot curve is empirical $i_w = Ae^{\gamma\rho^2} + Be^{-\gamma\rho^2}$.

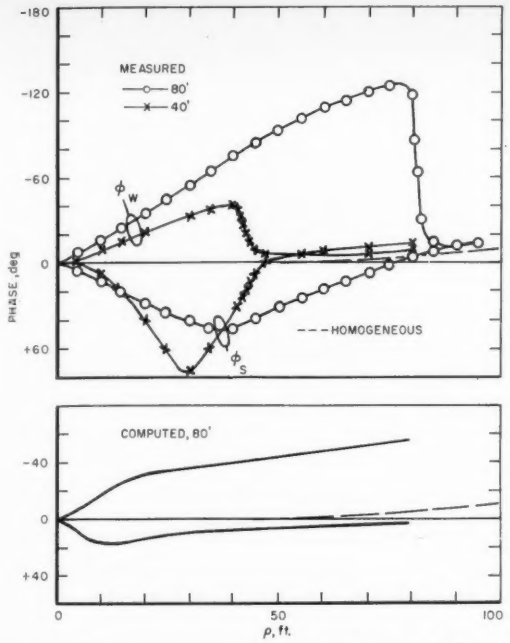


FIGURE 6. *Comparison of measured and theoretical phases for 40- and 80-ft 16-radial ground systems.*
Positive phase represents lead relative to downlead current.

the wires, where it showed a decided decrease. This one should expect from the interaction of fields of the parallel-flowing wire and soil currents, but the effect is quite localized.

With reference to the above, then, for $\rho < 20$ ft measured $|h_s|$ should be expected to fall below the computed values. Figures 4 and 5 show this to be true for both 40- and 80-ft systems. However, for $\rho > 20$ ft, the behavior of $|h_s|$ is distinctly different, the 40-ft system showing reduction in $|h_s|$ at all points, while the 80-ft system shows an enhancement. In the 40-ft system, the wire field in the first 10 ft is very nearly that predicted by current division according to parallel impedances, while for the 80-ft system agreement is only fair. In both systems, beyond 20 ft an increasing departure between measured and computed $|h_w|$ is evident, the nature of which appears different for the two situations.

It is tempting to say simply that, since beyond 20 ft the spacing s between the wires violates the condition $s << |\gamma_e^{-1}|$, they are acting no longer as a grid but as a collection of N independent, parallel, current filaments, each of whose internal impedance $Z_i = R_i + jX_i$ is given by the usual formulas [13]. Such a radial wire mesh would have at a distance ρ from the common center an equivalent surface impedance of $Z_w = 2\pi\rho Z_i/N$. At 40 ft, the sixteen No. 30 - copper - radial system would by this have $Z_w = 2.89/38.1°$. Since the soil impedance in parallel is $Z_s = 33.1/32.7°$ for $\sigma = 0.010$ and $k = 50$, eq (4) would yield $h_w = 0.0475$ m^{-1} for this distance; for ρ 60 ft, $h_w = 0.022$ m^{-1}. Experimentally, one finds in the 40-ft system $h_w = 0.055$ at 39 ft; in the 80-ft system, 0.052 and 0.0183 at 40 and 60 ft, respectively. Such agreement seems good until one examines the relation between observed and predicted phases, particularly in the light of the contrasting behavior of $|h_s|$, for the two radial lengths, in the region $\rho > 25$ ft. Thus, it would seem that the 80-ft system acts not merely as a grid of independent wires in parallel with the soil in the outer two thirds.

The nature of the departure of $|h_w|$ from the theoretical values is somewhat suggestive of standing waves. Assuming a transmission-line mode of behavior, one represents the current i_w in the wire at a distance x from some reference point by [14].

$$i_w = Ae^{\gamma_e x} + Be^{-\gamma_e x}, \tag{9}$$

where γ_e is the effective propagation constant of the medium surrounding the wire, and the ratio B/A is the reflection coefficient. Regarding the outer termination as the origin and looking back toward the downlead as the generator, the empirical fit with the experimental curve illustrated by the dash-dot curve of figure 5 is obtained for $\gamma = 0.10 + i0.197$, $A = 1$, $B = -\frac{1}{2}$, i.e., a lossy line terminated in an impedance higher than its characteristic impedance. Confirmation is had from the figures for σ and k derived from $\gamma_e^2 = (\gamma_0^2 + \gamma^2)/2$. These are $\gamma = 0.0064$ mho/m, and $k = 56.4$; also $\lambda = 32$ m, and $v = 0.16\ c$ where c is the velocity of light in a vacuum. The buried-transmission-line measurement gave respectively 0.010, 50, 30, and 0.15, which is a fairly reasonable agreement. Since in actuality the wires were a little below the interface, σ and k probably were between the above and the values 0.0033 and 29, respectively, obtained from $\gamma = \gamma_e$ (completely buried wire).

The observed phases seem to fit generally with the above picture. The change should be 180° in $\lambda/2 = 16$ m. The measured wire field phase referred to the downlead does not change by 180° in this distance, but advances more slowly at the rate of 1.90°/ft instead of 3.40°/ft. However, if at each point one considers the total phase change of wire field with reference to that at the same radius, the combined change is 1.86°+1.29°=3.15°/ft, out to about 50 ft where the total difference is about 130°. The return toward zero for the phase of the field between wires for $\rho > 50$ ft may be due to masking of the standing-wave circulating-current field in the soil by the incident downlead field. In the case of the 40-ft system, some evidence of the same thing shows up in the phase progression where in the soil it is 2.5°/ft, in the wire 1.3°/ft, total 3.8°/ft, but the magnitude curves, figure 4, show no recognizable standing wave characteristics. This may be because the region of independently acting radial length is considerably under half a wavelength (for the wire in the interface).

5. Conclusions

In both the 16-radial systems as well as in the 8-radial, it appears that a substantial potential difference existed between the wires and the soil. Hence, in no case studied experimentally can the grid and

soil be considered as surface impedances in simple parallel combination. However, in those regions where the ground-system parameters satisfy the remaining four assumptions made in the theory, the observed fields seem to agree with the predicted ones fairly well. In view of the similarity of soil conditions for the present experiment and the previous impedance studies on systems of many radials, it is likely that conclusions drawn here hold for those earlier studies as well.

To compute the H-field losses in a radial ground system laid on poorly conducting soil, one properly should consider four regions separately. The inner region is that in which the fields vary more rapidly than the limit specified in the theory; the second (in a proper design, by far the largest) is that in which at least the first four assumptions are satisfied; the third, that in which the wire spacing is such that the radials act independently; the fourth, the terminal region within a soil skin depth of the radial ends. In all four regions, copper losses would be correctly computed only from an observed current distribution, but it seems likely that an adequate approximation would result from using the distribution computed from the parallel impedance assumption; in any reasonable case these losses would be small. The ground losses for the inner region would have to be obtained from a volume current density distribution; they could be (usually are) eliminated by using an effectively solid sheet of highly conducting material in place of the radials. In the second region, where all assumptions but the fifth hold, the losses should be computable from Wait's theory even with very poor contact between wires and soil. The greater the number N, the better the results should be for losses computed in this second region. In the third and fourth regions, one has recourse to numerical integration of $\int Z_s H_s^2 dA$ where H_s $(\rho, 0)$ would have to be an experimentally determined quantity. For very poor soil, poor contact between radials and soil, and mismatched radial termination, sufficient standing-wave circulating current may exist to add a significant extra H-field loss component. Under these conditions it would seem worthwhile to direct attention toward providing a termination in the characteristic impedance of the radial regarded as a transmission line.

It must be remembered that the above-described conditions are also those for which considerable charging current power loss can appear in the contact impedance between radials and soil, as well as in the displacement flux terminating in the soil between radials, and in fact such "E-field" losses can be the dominating term in the total ground system loss picture, particularly where the loss tangent of the soil is of the order of unity. For better soils, these losses become of less importance, along with the standing-wave phenomenon, until for $\sigma/\omega k\epsilon_0 >> 10$ they likely are of little significance.

The authors express their appreciation to Harold Hanna and Anthony Muschbacher for their part in the construction, maintenance, and operation of the experimental facility under often very unfavorable climatic conditions, and to James R. Wait of NBS, Boulder Laboratories, for his many helpful comments and continued interest in the project.

6. References

[1] J. R. Wait and W. J. Surtees, Impedance of a top-loaded antenna of arbitrary length over a circular grounded system, J. Appl. Phys. **25,** 553 (1954).
[2] J. R. Wait and W. A. Pope, The characteristics of a vertical antenna with a radial-conductor ground system, Appl. Sci. Research **B4,** 177 (1954).
[3] J. R. Wait and W. A. Pope, Input resistance of l.f. unipole aerials, Wireless Engr **32,** 131 (1955).
[4] F. R. Abbott and C. J. Fisher, Design of ground system of radial conductors for a vlf transmitter, USNEL Rept. No. 105 (1949).
[5] H. A. Wheeler, Vlf antenna notebook-ground system-area losses by vertical electric field (unpublished memo, Wheeler Laboratories, Great Neck, N.Y.).
[6] J. R. Wait, On the calculation of transverse current loss in buried-wire ground systems, Appl. Sci. Research **B7,** 81 (1958).
[7] W. E. Gustafson, T. E. Devaney, and A. N. Smith, Model studies on vlf antenna ground system design, USNEL Rept. No. 822 (1957).
[8] T. E. Devaney, W. E. Gustafson, and N. H. Balli, Antenna model studies relating to the design of the North Atlantic vlf transmitting station, USNEL Rept. No. 884 (in preparation).
[9] W. G. Hutton, H. F. Cobb, and C. S. Warner, Navy vlf ground system design—field experimental data, 15 Jan 1957, and supplementary field experimental data 15 April 1957 (unpublished memos, Smith Electronics, Inc., Cleveland, Ohio).
[10] C. E. Smith, E. M. Johnson, W. G. Hutton, and G. F. Leydorf, Navy vlf ground system design, final report (submitted by Smith Electronics, Inc. to Development Engineering Corp., Washington, D. C., under prime contract NObsr–71360 of Continental Electronics Manufacturing Co., Dallas, Texas).
[11] J. R. Wait, A study of earth currents near a vlf monopole antenna with a radial wire ground system, Proc. IRE **46,** 1539 (1958).
[12] J. R. Wait, Earth currents near a top-loaded monopole antenna, J. Research NBS **62,** 247 (1959).
[13] American Institute of Physics handbook, pp. 5–91 and 5–201 (McGraw-Hill Book Co., Inc., New York, N.Y., 1957).
[14] J. C. Slater, Microwave transmission, first ed. p. 24 (McGraw-Hill Book Co., Inc., New York, N.Y., 1942).

San Diego, Calif. (Paper 63D2-16)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Diffraction of Electromagnetic Waves by Smooth Obstacles for Grazing Angles[1]

James R. Wait and Alyce M. Conda

(March 19, 1959)

The diffraction of electromagnetic waves by a convex cylindrical surface is considered. Attention is confined primarily to the region near the light-shadow boundary. The complex-integral representation for the field is utilized to obtain a correction to the Kirchhoff theory. Numerical results are presented which illustrate the influence of surface curvature and polarization on the diffraction pattern. Good agreement with the experimental results of Bachynski and Neugebauer is obtained. The effect of finite conductivity is also considered.

1. Introduction

The Kirchhoff theory of diffraction has been quite successful in the prediction of field strength when the transmission path is obstructed by obstacles such as mountainous ridges [1–8].[2] In its usual form, however, this theory disregards the difference between vertical and horizontal polarization. Furthermore, the electrical properties of the diffracting obstacle are not considered since in effect, the obstacle is regarded as a "black screen" or absorbing knife edge. A brief mention of the various developments, both old and new, will be made in order to properly orient the reader.

Schelleng, Burrows, and Ferrell [1] in a now classic paper, applied knife edge Kirchhoff theory to the computation of the fields behind mountainous obstacles and other pronounced terrain features. The influence of the terrain in the immediate vicinity of the terminals was accounted for by the introduction of ground-reflected rays which are also diffracted by the obstacles.[3] A similar approach was made by Dickson, Egli, Herbstreit, and Wickizer [5], who demonstrated that the presence of an obstacle can, in some cases, produce an increase of field strength beyond that expected in the absence of the obstacle. This phenomenon has been described rather aptly as "obstacle gain" and has been used to advantage in the siting of the terminal and repeater stations in microwave relay links.

The Japanese workers [4,6] have also been very active in this field. Presumably, the mountainous features of their country have motivated much of this work. Matsuo [4] in particular, has applied the Kirchhoff theory to many geometrical forms including conical shaped obstacles. Some fundamental theoretical work has been carried out by Furutsu [9] who has presented solutions for diffraction by very general forms of obstacles.

S. O. Rice [11] has given a comprehensive solution for the diffraction of a plane wave by a perfectly conducting parabolic cylinder which was considered to be a more realistic model of a terrain obstacle. Particular attention was paid to the fields near the horizon plane on the boundary region between light and shadow. It was shown that to a first approximation, the diffraction pattern behind the obstacle was very similar to that for a knife edge. The main effect of the finite radius of curvature of the crest was to increase the field strength for vertical polarization and to decrease it for horizontal polarization. A similar observation had been made by Artmann [12] sometime earlier who employed a circular cylindrical model.

Still another approach to this problem was adopted by Norton, Rice, and Vogler [13]. They represented the diffracting obstacle first by a sphere for the case of a very large radius of curvature and then by a knife edge for vanishing small radii of curvature. A curve of diffraction loss as a function of radius of curvature was then drawn by interpolating between these two limiting cases. According to their results, the diminishing of the radii of curvature is ac-

[1] A preliminary version of this paper was presented at the International Conference on Wave Propagation, Liège, Belgium on October 8, 1958.

[2] Figures in brackets indicate the literature references at the end of this paper.

[3] A scheme which will suppress or modify these ground-reflected rays has been suggested by H. E. Bussey (Reflected Ray Suppressions, Proc. IRE **38**, 1453, 1950).

508098—59——5

companied by enormous increases in the field strength even though the separation between terminals and the angular distance is unchanged.

Refinements of Kirchhoff diffraction theory when applied to the diffraction by smooth cylindrical models were made by Neugebauer and Bachynski [14, 15]. The basic idea of their analysis is that the equivalent radiating aperture is illuminated not only by the incident rays, but from rays reflected from the illuminated side of the cylindrical crest. The subsequent radiation from this aperture plane, which extends from the apex of the crest upward to infinity, is again reflected from the shadow side of the crest. Neugebauer and Bachynski are quite willing to admit that this is not a rigorous procedure, but nevertheless, it leads to results which are in good agreement with experiment.

It is the purpose of the present paper to present a method for the computation of diffracted fields by convex surfaces which is both rigorous and easy to apply. The mathematical basis of this method was described in a previous paper [16]. This was an extension of the fundamental work of Van der Pol and Bremmer, and Fock. The reader is referred to the above quoted paper for an account of the early theoretical work.

Unfortunately, the existing solutions for the diffraction by spheres and cylinders are poorly convergent in the line-of-sight region. Consequently, these formal solutions require a great deal of numerical work despite the fact that they are mathematically elegant. A significant break-through was made when it was suggested that the complex-integral representation of the field was easier to evaluate by numerical means than attempting to sum the residues of the poles of the integrand. While this point was probably recognized by Van der Pol and Bremmer [17], in the 1930's, it was not until 1946 that Academician V. A. Fock actually demonstrated this in what is now regarded as a classic paper [18]. Fock's method was applied by the present authors to the computation of the radiation pattern of an antenna mounted on or near a convex lossy surface [16]. The present problem is really an extension of this work.

2. Formulation

The problem is to calculate the field at some point P exterior to the cylinder in terms of the field on its surface (see fig. 1a). Since the incident magnetic field H^{inc} (ρ, ϕ) is taken to have only a z component, the resultant field H (ρ, ϕ) also has only a z component. An immediate application of Green's theorem yields

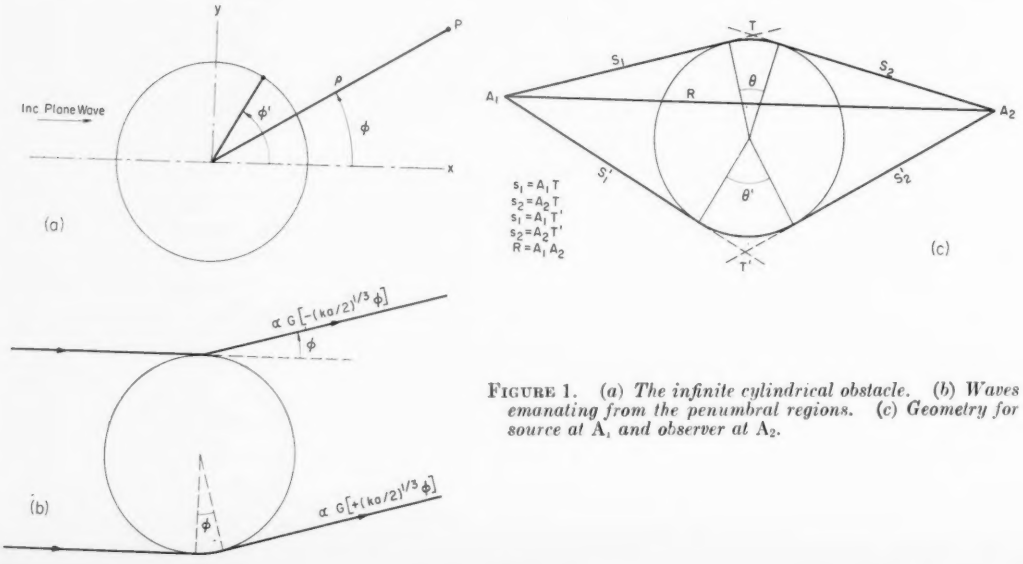


FIGURE 1. *(a) The infinite cylindrical obstacle. (b) Waves emanating from the penumbral regions. (c) Geometry for source at* A_1 *and observer at* A_2.

$$H_z(\rho,\phi)=H_z^{\text{inc}}(\rho,\phi)+a\int_0^{2\pi}\left[H_z(\rho',\phi')\frac{\partial}{\partial\rho'}Q(\rho,\phi;\rho',\phi')-Q(\rho,\phi;\rho',\phi')\left(\frac{\partial H_z(\rho',\phi')}{\partial\rho'}\right)\right]_{\rho'=a}d\phi', \tag{1}$$

where the Green's function is given by

$$Q(\rho,\phi;\rho',\phi')=-\frac{i}{4}H_0^{(2)}(k\hat{\rho}), \tag{2}$$

with

$$\hat{\rho}=[\rho^2+(\rho')^2-2\rho\rho'\cos(\phi-\phi')]^{1/2}.$$

The primed coordinates ρ', ϕ', with $\rho'=a$, refer to a variable point on the surface of the cylinder whereas ρ, ϕ, of course, refer to the position of the observer. The above integral formula can be simplified somewhat when the Leontovich boundary condition can be used [16]. Then

$$\frac{\partial H_z(\rho',\phi')}{\partial\rho'}\Big]_{\rho'=a}=i\epsilon_0\omega ZH_z(a,\phi), \tag{3}$$

where Z is the surface impedance of the cylinder and ϵ_0 is the permittivity of the surrounding space.[4] Invoking this approximation, the external field is expressed explicitly in terms of the surface field $H_z(a,\phi)$ which can be regarded as known. The evaluation of the integral, however, is not readily carried out unless certain restrictions are made. For the moment it will be assumed that the observer is at a large distance from the cylinder. Consequently,

$$\begin{aligned}Q&\cong-\frac{i}{4}\frac{1}{\sqrt{k\hat{\rho}}}e^{i[(\pi/4)-k\hat{\rho}]}\\&\cong\frac{e^{-i\pi/4}}{4\sqrt{k\rho}}e^{-ik\rho}e^{ik\rho'\cos(\phi-\phi')},\end{aligned} \tag{4}$$

where only the first term in the asymptotic expansion for $H_0^{(2)}(k\hat{\rho})$ has been retained.

In a previous paper [16], it was shown that

$$\begin{aligned}H_z(a,\phi')=H_0\sum_{m=0}^{\infty}e^{-ika\left[\left(\frac{\pi}{2}-\phi'\right)+2\pi m\right]}g\left[\frac{\pi}{2}-\phi'+2\pi m\left(\frac{ka}{2}\right)^{1/3}\right]\\+H_0\sum_{m=0}^{\infty}e^{-ika\left[\left(\phi'+\frac{\pi}{2}\right)+2\pi m\right]}g\left[\phi'+\frac{\pi}{2}+2\pi m\left(\frac{ka}{2}\right)^{1/3}\right],\end{aligned} \tag{5}$$

for the region $0\leqq|\phi'|\leqq(\pi/2)+\Delta\phi$, where $(\Delta\phi)^3<<1$. The function $g(X)$ occurring in the above expression is given by [16]

$$g(X)=\frac{1}{\sqrt{\pi}}\int_{\infty e^{-i2\pi/3}}^{\infty}\frac{e^{-iXt}}{w_1'(t)-qw_1(t)}dt, \tag{6}$$

where $w_1(t)$ and $w_1'(t)$ are Airy integrals defined in the appendix, and

$$q=-i\left(\frac{ka}{2}\right)^{1/3}Z/\eta_0,\quad\text{with }\eta_0\cong120\pi.$$

The quantity m can be physically regarded as the number of times a wave has crept around the cylinder. For $(ka)^{1/3}\pi>>1$, only the term $m=0$ need be retained. Furthermore, if $ka>>1$ and $|\phi|$ less and not near $\pi/2$ the geometrical optical approximation for $H_z(a,\phi')$ is applicable. This reads

$$H_z(a,\phi')\cong H_0(1+R_e e^{-ika\cos\phi'}), \tag{7}$$

[4] Equation (3) is strictly valid only when $(Z)<<120\pi$ but actually this is more stringent than necessary [16].

where R_v is a Fresnel reflection coefficient given by

$$R_v=-\frac{(Z/\eta_0)-\cos\phi'}{(Z/\eta_0)+\cos\phi'}. \tag{8}$$

3. Evaluation of the Field

The secondary or scattered field can thus be written in the following form

$$H_z^s(\rho,\phi)=\frac{e^{-i(k\rho-\pi/4)}}{2\sqrt{2\pi k\rho}}(\psi_1+\psi_2), \tag{9a}$$

where

$$\psi_1=\int_{-\pi}^{\infty}[\cos(\phi'-\phi)-kaZ/\eta_0]e^{-ika[(\pi/2)+\phi'-\cos(\phi'-\phi)]}g(X_1)d\phi' \tag{9b}$$

$$\psi_2=\int_{-\infty}^{\pi}[\cos(\phi'-\phi)-kaZ/\eta_0]\,e^{-ika[(\pi/2)-\phi'-\cos(\phi'-\phi)]}g(X_2)d\phi'. \tag{9c}$$

The phase in the integral ψ_1 has the form

$$\Omega(\phi')=ka[(\pi/2)+\phi'-\cos(\phi'-\phi)],$$

so that the stationary points are at $\phi'=\phi_n'=(4n-1)\ \pi/2+\phi$ for $n=0,\ 1,\ 2,\ 3\ \ldots$. The phase function is expanded about ϕ_n' and is of the form

$$\Omega(\phi')=ka[2\pi n+\phi+\tfrac{1}{6}(\phi'-\phi_n')^3],$$

neglecting terms containing higher power than the third. The integrals over ϕ' are now in the form of Airy integrals and thus ψ_1 and ψ_2 can be expressed in the form

$$\psi_1=\frac{-2\sqrt{\pi}e^{-i\pi/4}}{(ka/2)^{2/3}}\,e^{-ika\phi}\sum_{m=0}^{\infty}e^{-i2\pi mka}\,\hat{G}\left[\left(\frac{ka}{2}\right)^{1/3}\phi\right], \tag{10a}$$

$$\psi_2=\frac{-2\sqrt{\pi}e^{-i\pi/4}}{(ka/2)^{2/3}}\,e^{-ika\phi}\sum_{m=0}^{\infty}e^{-i2\pi mka}\,\hat{G}\left[-\left(\frac{ka}{2}\right)^{1/3}\phi\right], \tag{10b}$$

where

$$\hat{G}(X)=\frac{e^{-i\pi/4}}{\sqrt{\pi}}\int_{\infty\,e^{-i2\pi/3}}^{\infty}e^{-iXt}\frac{v'(t)-qv(t)}{w_1'(t)-qw_1(t)}\,dt, \tag{11}$$

where $w_1(t)$ and $v(t)$ are Airy integrals defined in the appendix. The quantity t is a new variable of integration.

The special cases corresponding to $q=0$ and $q=\infty$ lead directly to Goriainov [19] who assumed at the outset that the cylinder was perfectly conducting. These special cases are discussed in the subsequent text. The total field scattered by the cylinder is thus expressed by

$$H^s=-H_0\frac{e^{-ik\rho}}{\sqrt{k\rho}}(h_0+h_1), \tag{12}$$

where

$$h_0=\sqrt{\frac{2}{\pi}}\,e^{i\pi/4}\frac{\sin ka\phi}{\phi}, \tag{13}$$

and

$$h_1=+\sqrt{2}\left(\frac{ka}{2}\right)^{1/3}\{e^{-ika\phi}G[(ka/2)^{1/3}\phi]+e^{ika\phi}G[-(ka/2)^{1/3}\phi]+e^{-ika(2\pi+\phi)}G[(ka/2)^{1/3}(2\pi+\phi)]$$
$$+e^{-ika(2\pi-\phi)}G[(ka/2)^{1/3}(2\pi-\phi)]+\ \ldots\ ]\},\quad (14a)$$

where

$$G(X)=\hat{G}(X)+\frac{e^{-i\pi/4}}{2X\sqrt{\pi}}. \quad (14b)$$

The preceding results will now be generalized to include the case of a spherical wave incidence. The source is now imagined to be a magnetic dipole, located at $(\rho_0, \pi, 0)$, thus the radius of curvature of the incident wave at the origin is ρ_0. Equation (12) for the scattered field in the transverse plane $z=0$ is now simply generalized to spherical wave incidence by multiplying it by $\sqrt{\rho_0/(\rho+\rho_0)}$. This is valid for any form of cylindrical obstacle subject to ρ and ρ_0 both large compared to the maximum transverse dimension of the effective scattering region.[5] Thus

$$H_z^s\cong -H_0\frac{e^{-ik\rho}}{(k\rho)^{1/2}}\left(\frac{\rho_0}{\rho_0+\rho}\right)^{1/2}(h_0+h_1), \quad (15)$$

where H_0 is the value of the incident field reckoned at the center of the cylinder and h_0 and h_1 are as defined above.

The portion of the scattered field proportional to h_0 corresponds to Kirchhoff diffraction since it can be readily predicted by physical optics. It is independent of the electrical properties of the obstacle (and thus independent of the surface impedance Z) and is proportional to the cross-sectional dimension of the obstacle for small deflection angles. The remaining part of the scattered field, proportional to h_1, can be regarded as a correction to Kirchhoff theory. When $ka\to\infty$, the ratio h_1/h_0 asymptotically approaches zero. For finite but large values of ka and small values of ϕ only the first two terms containing $G[(ka/2)^{1/3}\phi]$ and $G[-(ka/2)^{1/3}\phi]$ are significant. These can be interpreted as waves emanating from the penumbral regions on the cylinder as indicated in the sketch shown in figure 1b. In this case, the distances ρ and ρ_0 are both assumed very large compared to the transverse dimension of the obstacle (i.e., $2a$). To relax this condition, it is necessary to return to the original integral for the field and use the following far-field approximation for the Green's function:

$$Q\cong -\frac{i}{4}\frac{1}{\sqrt{k\hat{\rho}}}\,e^{i(\pi/4-k\hat{\rho})}, \quad (16)$$

where $\hat{\rho}$ is not assumed to be large compared to $2a$ but is still large compared to the wavelength illustrated in figure 1c. Carrying through the analysis leads to

$$H_z=H_z^p+H_z^s\cong\frac{e^{-ikR}}{R}\,[F(X,u)+F(X',u')], \quad (17)$$

with

$$F(X,u)=\frac{e^{i\pi/4}}{\sqrt{\pi}}\int_a^\infty e^{-i\alpha^2}d\alpha-\frac{G(X)}{u}\,e^{-iu^2}, \quad (18)$$

where

$$u=\frac{\alpha}{X}=\left(\frac{2ks_1s_2}{s_1+s_2}\right)^{1/2}\frac{(2/ka)^{1/3}}{2},$$

and

$$X=(ka/2)^{1/3}\theta.$$

Equation (18) is only valid when u is somewhat greater than unity. A more accurate form is obtained when $G(X)$ is replaced by $G(X)-\frac{i}{4u^2}\frac{\partial^2 G(X)}{\partial^2 X}$.

[5] See J. R. Wait, Appl. Sci. Research **B.4,** 464 (1955) for example.

The function $F(X', u')$ is the same form as $F(X, u)$, and u' and X' are defined by

$$u'=\frac{\alpha'}{X'}=\left(\frac{2ks_1's_2'}{s_1'+s_2'}\right)^{1/2}\frac{(2/ka)^{1/3}}{2},$$

and

$$X'=(ka/2)^{1/3}\theta'.$$

The quantities s_1 and s_2 are the linear distances from the source A_1 and the receiver A_2 measured to the intersection of tangent planes of the surface of the cylinder as sketched in figure 1c. Similar remarks apply to the primed quantities.[6] The quantity R is the linear distance separating A_1 and A_2.

It should be noted that when A_1 and A_2 recede to infinity, the earlier (Fraunhofer) solution is retrieved as a special case. To effect this transition, it is convenient to use the asymptotic relation

$$[uF(X, u)]\cong e^{-i\alpha^2}\left[\frac{e^{-i\pi/4}}{2\sqrt{\pi}X}-G(X)\right] \quad \text{for large positive } \alpha \text{ and} \tag{19}$$

$$\cong e^{-i\alpha^2}\left[u+\frac{e^{-i\pi/4}}{2\sqrt{\pi}X}-G(X)\right] \quad \text{for large negative } \alpha.$$

Up to this point the analysis has been developed explicitly for the incident magnetic field vector directed along the axis of the cylinder. The results can be transformed to the case for the incident electric field vector directed along the axis of the cylinder by making the following transformation

$$H_z^{\text{inc}}(\rho, \phi)\rightarrow E_z^{\text{inc}}(\rho, \phi)$$

$$H_z^s\ (\rho, \phi)\rightarrow E_z^s\ (\rho, \phi)$$

$$\left[q=-i\left(\frac{ka}{2}\right)^{1/3}\left(\frac{Z}{\eta_0}\right)\right]\rightarrow\left[q=-i\left(\frac{ka}{2}\right)^{1/3}\left(\frac{\eta_0}{Z}\right)\right].$$

For metallic surfaces, $q\cong 0$ for H parallel polarization and $q\cong\infty$ for E parallel polarization, and these limiting cases are referred to in the following section as vertical and horizontal polarizations, respectively.

4. Numerical Results

Basic to the present work is the function $G\ (X)$. In the region near $X=0$, the relevant integrals cannot apparently be evaluated in closed form. The integrands are quite well behaved however, and it is not too difficult to evaluate them by numerical means. In figures 2a and 2b the real and imaginary parts of $G\ (X)$ are plotted as a function of X for $q=0$ and $q=\infty$, respectively.

Before discussing the diffraction pattern, it is of interest to consider the behavior of the field along the horizon plane. Clearly, this corresponds to $\theta=0$, $X=0$, and $\alpha=0$. The function $F\ (\alpha)$ then becomes

$$F\ (0)=\frac{1}{2}-\frac{G(0)}{u}, \tag{20}$$

where

$$u=\left(\frac{2ks_1s_2}{s_1+s_2}\right)^{\frac{1}{2}}\frac{(2/ka)^{\frac{1}{3}}}{2}, \text{ as before.}$$

Numerical integration yields the following values: for $q=0$, $G\ (0)=-0.295+i\ 0.0811$; and

[6] θ and θ' are the angles subtended by the tangent planes as indicated in figure 1c. They are referred to sometimes as angular distances and expressed in milliradians.

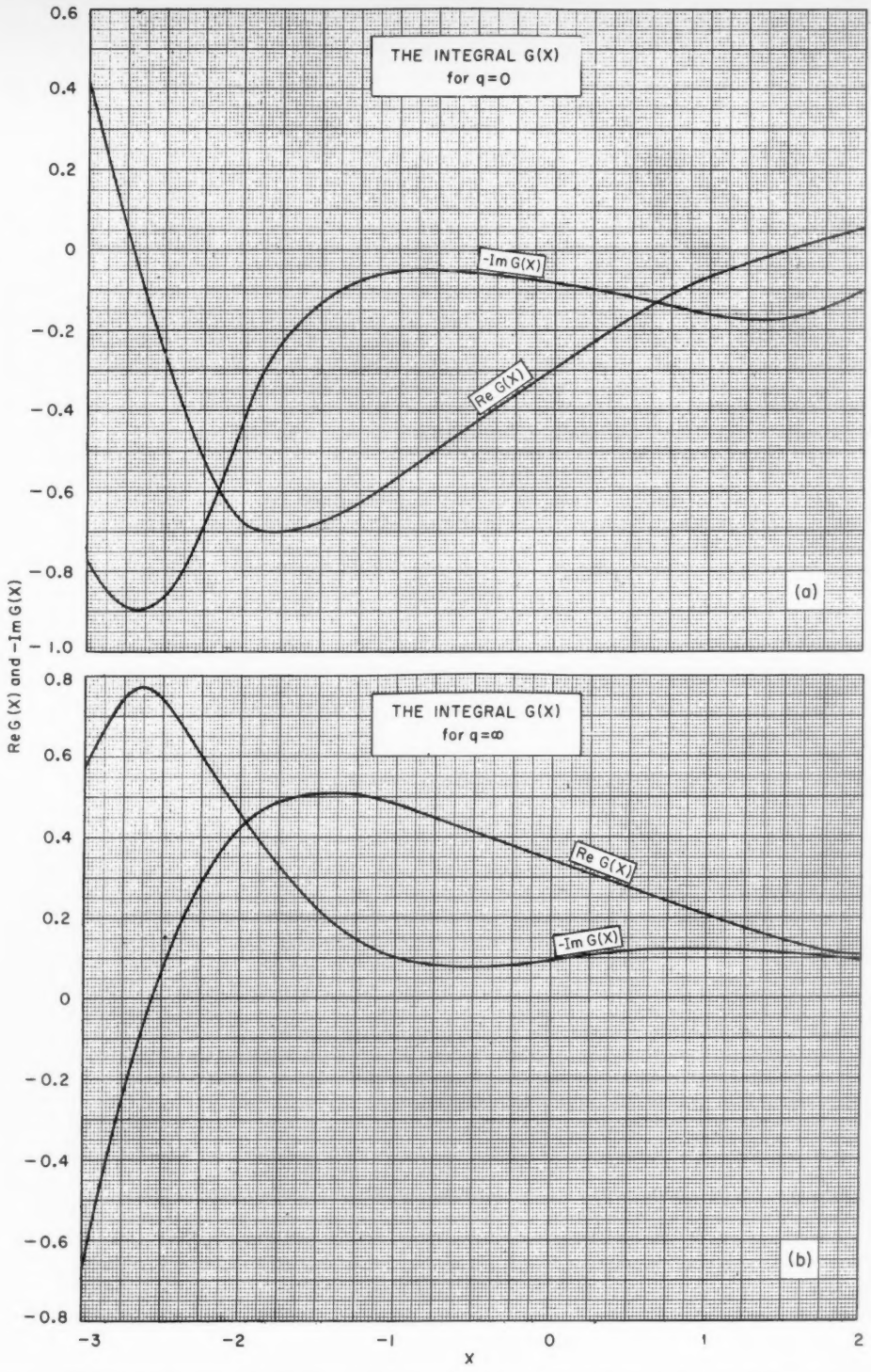


FIGURE 2. *(a) Real and imaginary parts of the integral* G(X) *as a function of* X *for* q = 0. *(b) Real and imaginary parts of the integral* G(X) *as a function of* X *for* q = ∞.

for $q=\infty$, $G(0)=0.344-i\,0.0918$. This indicates that for vertical polarization (corresponding to $q=0$) with a metallic convex surface, the field along the shadow boundary is increased above that for a knife edge. The converse is true for horizontal polarization (corresponding to $q=\infty$). The increase or reduction from ½ is approximately the real part of the function $[-G(0)/u]$. Therefore for $q=0$, $|F(0)|\cong 0.500+(0.295/u)$ and for $q=\infty$, $|F(0)|\cong 0.500-(0.344/u)$. It is of interest to compare this departure from ½ with Rice's prediction using a parabolic cylinder model. The agreement is good if we equate the curvature of the apex of his parabolic cylinder with our a. It is also noted that his assumed plane-wave incidence is equivalent to moving our source to infinity (i.e., $d_1=\infty$). Then on identifying Rice's h with $ka/2$ and his Λ with $[(s_1+s_2)/(2\pi k s_1 s_2)]^{\frac{1}{2}}$, we find according to Rice for $q=0$, $|F(0)|\cong 0.500+(0.296/u)$ and for $q=\infty$, $|F(0)|\cong 0.500-(0.342/u)$, being in close agreement with our values quoted above. On the other hand, Artmann, using a circular cylindrical model, gives for $q=0$, $|F(0)|\cong 0.500+(0.165/u)$ and for $q=\infty$, $|F(0)|\cong 0.500-(0.33/u)$. The disagreement in the case for $q=0$ with our result should be noted. It is suggested that the asymptotic approximation for the Hankel functions used by Artmann is not fully justified.

Neugebauer and Bachyinski have also computed a value on the shadow boundary using a semiempirical modification to Kirchhoff theory. Again transferring to our notation they would give[7]

$$\text{for } q=0, \qquad |F(0)|\cong 0.500+\frac{0.271}{u}, \quad \text{and}$$

$$\text{for } q=\infty, \qquad |F(0)|\cong 0.500-\frac{0.282}{u}.$$

The agreement of these values with ours (and those of Rice) is reasonable in view of the entirely different theoretical approaches employed. It is probable that at least part of the discrepancies can be attributed to the nonvalidity of some of the geometrical-optical arguments used by Neugebauer and Bachynski.

5. Diffraction Patterns

The diffraction pattern behind the obstacle is specified by the behavior of the function $F(X, u)$ as α or uX varies. The amplitude $|F(X, u)|$ itself and the resulting diffraction loss L_a are plotted in figure 3a for various values of u with $q=0$ and ∞. As u tends to infinity, the patterns approach the Kirchoff form described by the function $F_0(\alpha)$. Since the function $F_0(\alpha)$ occurs frequently in what follows, its real and imaginary parts are plotted in figure 3b.

The diffraction pattern behind a metallic cylindrical model of a mountain crest has been measured by Neugebauer and Bachynski in the K-band frequency range. For purposes of comparison, they also duplicated the experiment using a thin metallic sheet which was essentially equivalent to a knife edge. Their results for $ka=239$ are shown in figure 4a and 4b for vertical polarization ($q=0$) and horizontal polarization ($q=\infty$), respectively. The ordinates are relative power and the abscissa are distance d_2 measured from the crest to the observer. The quantity h is the vertical displacement of the receiver above or below the horizon plane. In both figures 4a and 4b the source which is a transmitting horn is at a fixed location, being a distance d_1 ($=150\lambda$) from the crest. Values calculated from eq (18) are also shown on the curves. The agreement is reasonably good.

Neugebauer and Bachynski also measured the fields behind the crest for various values of the curvature parameter ka. This is shown in figure 5 where the ordinate is relative power and the abscissa is the deflection angle θ. The corresponding calculated curves are also shown in figure 5, taking $d_1=150\lambda$ and $d_2=113\lambda$. The agreement is still fairly good and probably is within experimental error. These curves illustrate in a striking way the influence of curvature on the diffraction fields. In the case of horizontal polarization and deep in the shadow, the

[7] Their quantity 'A'$=-8\pi^{1/2}2^{-1/3}$ times the real part of our $G(0)$.

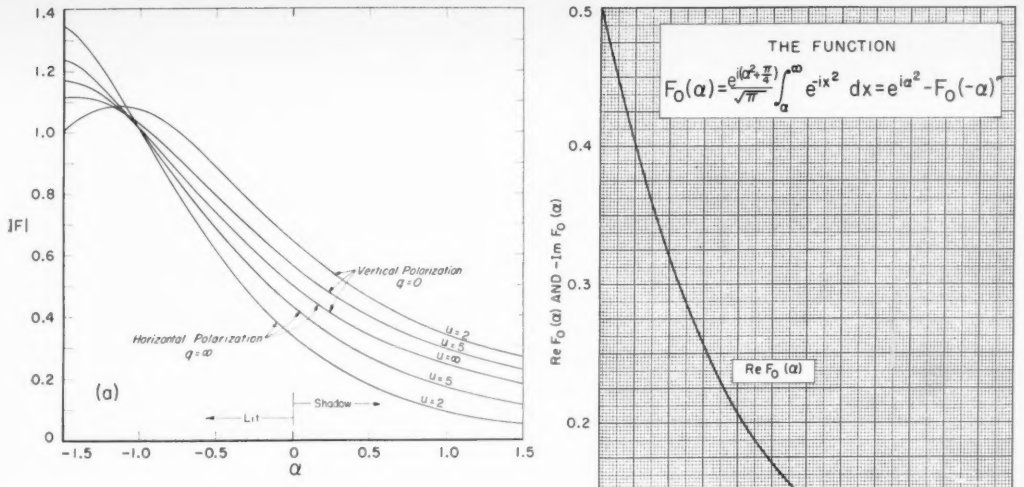


FIGURE 3. *(a) Diffraction pattern for a conducting convex surface. (b) Real and imaginary parts of Fresnel integral* $F_0(\alpha)$ *as a function of* α.

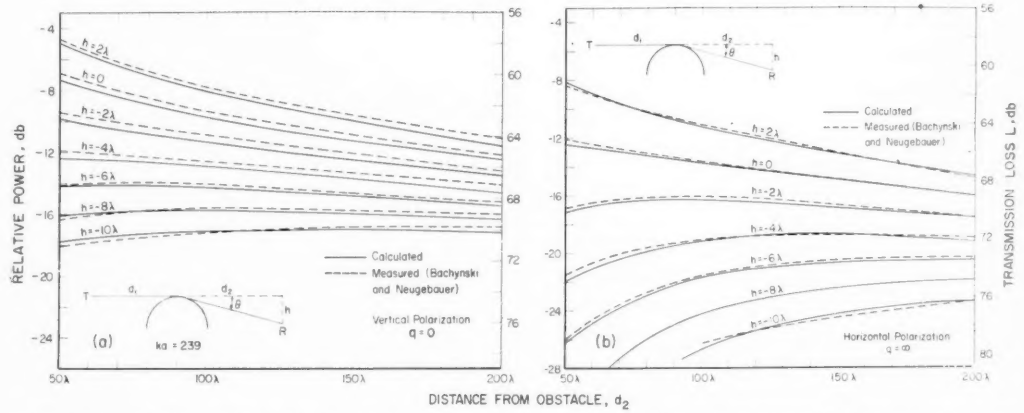


FIGURE 4. *(a) Power distribution behind a cylindrical obstacle. (b) Power distribution behind a cylindrical obstacle.*

finite radius of curvature greatly reduces the fields below the Kirchhoff or knife-edge value. This is in accord with the results of Norton, Rice, and Vogler [13].

6. Further Numerical Results

To obtain results for finite conductivity, it is necessary to evaluate the relevant integrals for finite values of q. For the case of H-parallel or vertical polarization

$$q=-i\left(\frac{ka}{2}\right)^{1/3}Z/\eta_0,$$

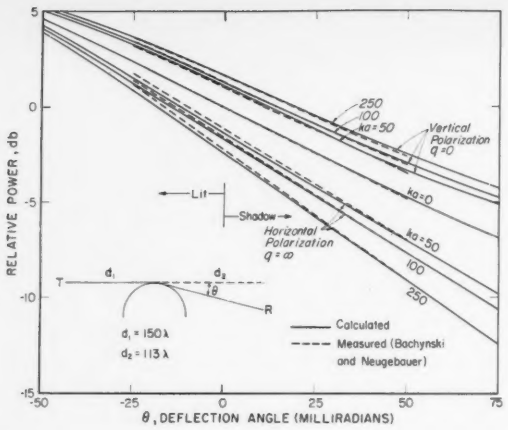


FIGURE 5. *Effect of curvature on diffraction by a convex surface.*

where Z is the surface impedance. Now if the surface is a good conductor

$$Z \cong (i\mu\omega/\sigma)^{1/2},$$

subject to $\sigma >> \epsilon\omega$. That is, displacement currents in the conductor are neglected. The consequence of this approximation was discussed in some detail in a previous paper devoted to this subject [16]. Following the earlier notation

$$q \cong |q| e^{-i\pi/4} = 2^{1/6} A e^{-i\pi/4},$$

where

$$A = (ka)^{1/3} \left(\frac{\epsilon_0 \omega}{2\sigma}\right)^{1/2},$$

is the "loss factor" for the surface. It is seen that for infinite conductivity, A is zero. On the other hand, for very poor conductivity and large values of ka, the "loss factor", A, may approach infinity. These two limiting cases correspond to $q=0$ and $q=\infty$, respectively, and were discussed in the previous section. More usually A takes some intermediate value.

For the general case

$$\hat{G}(X) = \hat{G}_1(X) + \hat{G}_2(X) \tag{21}$$

$$G(X) = \hat{G}_1(X) + \hat{G}_3(X), \tag{22}$$

where

$$\hat{G}_1(X) = \frac{e^{-i\pi/4}}{\sqrt{\pi}} \int_0^\infty e^{-iXt} \frac{v'(t) - qv(t)}{w_1'(t) - qw_1(t)} dt, \tag{23}$$

$$\hat{G}_2(X) = \frac{e^{i3\pi/4}}{2X\sqrt{\pi}} + \hat{G}_3(X), \tag{24}$$

$$\hat{G}_3(X) = \frac{e^{i\pi/12}}{\sqrt{\pi}} \int_0^\infty e^{(-Xt/2)(\sqrt{3}-i)} \left[\frac{v'(t) - qe^{-2\pi i/3} v(t)}{w_2'(t) - qe^{-2\pi i/3} w_2(t)}\right] dt. \tag{25}$$

The above relations follow immediately from eq (11) and (14b). The real and imaginary parts of the integrals $\hat{G}_1(X)$ and $\hat{G}_3(X)$ were evaluated numerically using tabulated values of the Airy integral functions. For the present purpose only the real and imaginary parts of the function $G(X)$ are required.

As mentioned above, the field along the horizon plane is of special interest. Expressed as a ratio to the free-space field, this can be written

$$F(0)=\frac{1}{2}-\frac{G(0)}{u}.$$

The real and imaginary parts of G (0) are plotted in figure 6 in the form of an Argand diagram. Various values of A are shown on the curve. The left-hand portion of the curve (i.e., smaller A values) corresponds to an increase over the knife edge value. On the other hand the right-hand portion (i.e., larger A) corresponds to a decrease below the knife-edge field. It is of interest to note that for an A value of about 0.4 the magnitude of the diffraction field is almost the same as that of the knife-edge field. Associated with these intermediate values of A, however, are appreciable phase shifts which correspond to a lag relative to the phase of the knife-edge field. The end points of the curve corresponding to $q=0$ and $q=\infty$ were discussed in the earlier section.

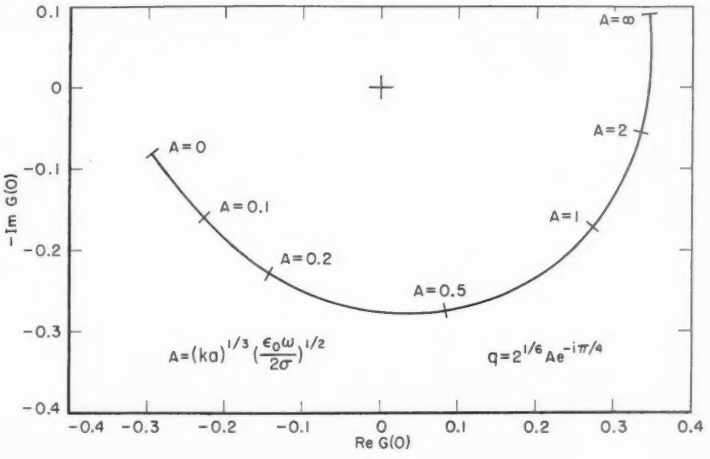


FIGURE 6. *Integral* G(O) *plotted in the complex plane for various values of the conductivity parameter* A.

Applicable to vertical polarization.

The real and imaginary parts of G (X) are shown plotted in figures 7a and 7b, respectively, as a function of X for various values of A.

The diffraction pattern defined by $|F(X, u)|$ can thus be readily computed for any value of A. Examples of these are shown in figures 8a, 8b, and 8c for u equal to 2, 5, and 10, respectively.

The preceding results indicate that the influence of finite conductivity can be marked for H-parallel or vertical polarization. For purposes of illustration, the displacement currents were neglected so that q has a phase angle of -45 deg. For poorer conductivities this restriction may be violated and the numerical results are not strictly applicable although they can be used for approximate estimates as long as the conduction currents are at least greater than the displacement currents. (i.e., $\sigma > \epsilon\,\omega$). In this case, the equation for A should be replaced by

$$A=(ka)^{1/3}\left(\frac{\epsilon_0\omega}{\sqrt{\sigma^2+\epsilon^2\omega^2}}\right)^{1/2}\frac{1}{2^{1/2}}.$$

The influence of finite conductivity for E-parallel or horizontal polarization is negligible for most cases since $|q|$ is always very large compared to unity.

7. Concluding Remarks

A theory for diffraction of electromagnetic waves by a smooth and convex lossy obstacle has been presented. The model employed for computation is a circular cylinder of finite conductivity.

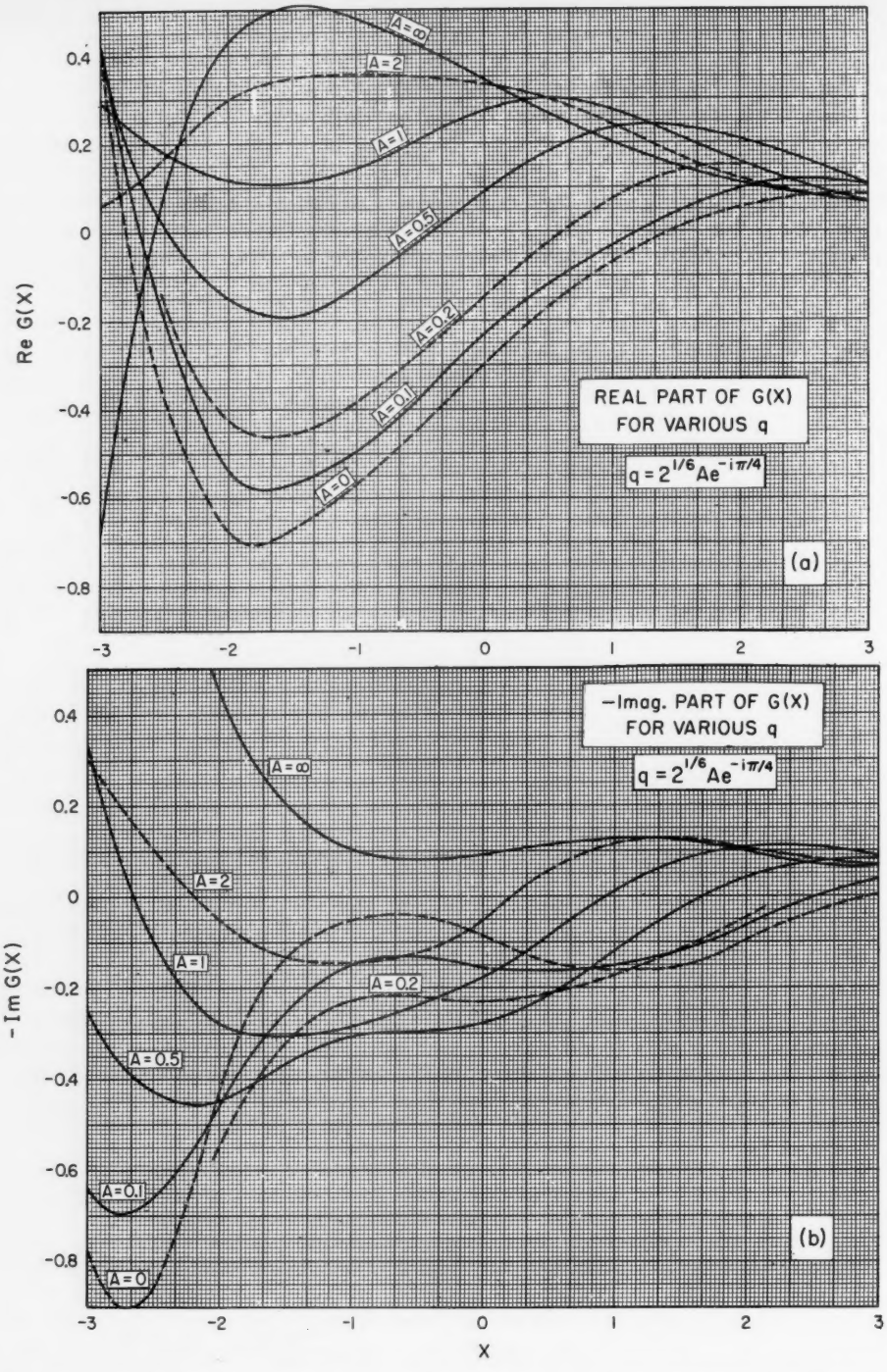


FIGURE 7. *(a) The real part of the function* G(X) *as a function of* X *for various values of the conductivity parameter A. (b) The imaginary part of the function* G(X) *as a function of* X *for various values of the conductivity parameter A.*

Both (a) and (b), applicable to vertical polarization.

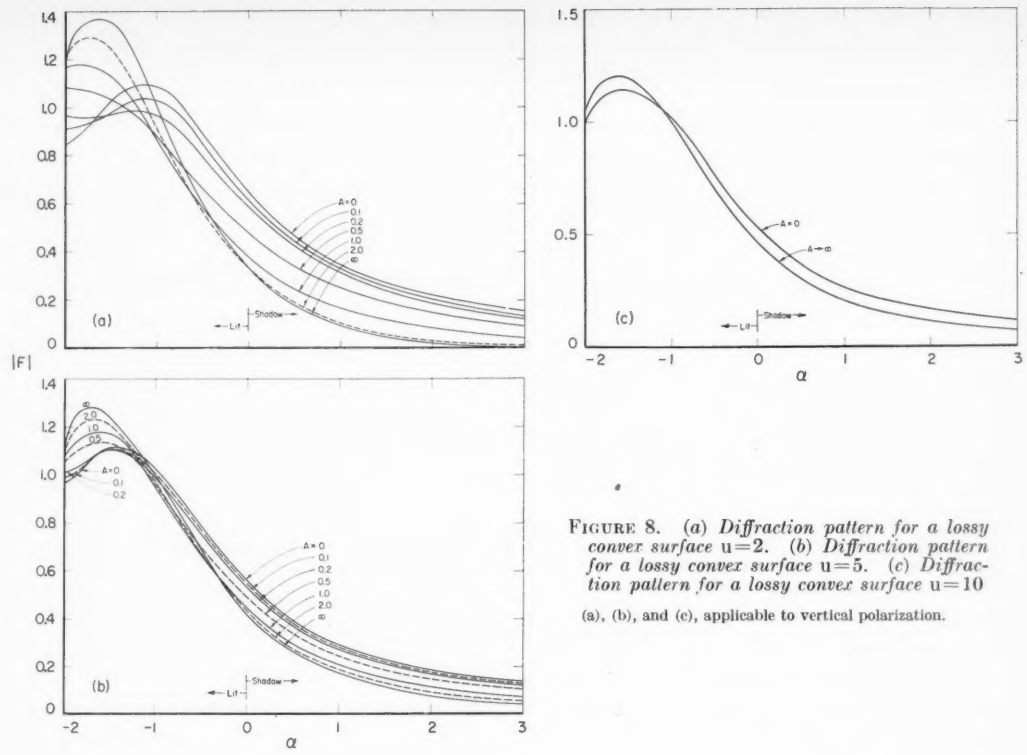


FIGURE 8. *(a) Diffraction pattern for a lossy convex surface* u=2. *(b) Diffraction pattern for a lossy convex surface* u=5. *(c) Diffraction pattern for a lossy convex surface* u=10

(a), (b), and (c), applicable to vertical polarization.

Probably the most important conclusion is that the influence of finite curvature of the diffracting metallic obstacle is to add a positive correction to the Kirchhoff computed field for a vertically polarized incident field, whereas the correction is negative for horizontally polarized incident field. The increase or decrease of amplitude from the Kirchhoff result is accompanied by a corresponding change of phase which may be important in certain applications.

The influence of finite conductivity is appreciable for vertical polarization and the curves presented in the paper may be used to compute this effect. In general, the ohmic loss tends to reduce the resultant diffracted field. In the case of a horizontal polarization, the finite conductivity has a negligible effect and the curves presented in the present paper for $q=\infty$ may be used directly.

While the present results are formally restricted to a diffracting surface of constant curvature, it can be expected that the results can be used for generally convex surfaces if the curvature is slowly changing. For example, it may be noted that the diffraction pattern has the same approximate analytical form for a circular cylinder, a sphere and a parabolic cylinder. At least this is true if the radius of curvature is large compared to the wavelength and the deflection angles are small. Deep in the shadow zone, this similarity principle breaks down and the field depends on the rate of change of the curvature. This aspect of the problem has been discussed recently by Wu [24] who has generalized some of the quasi-geometrical concepts of Keller [25]. As Wu points out, however, for small deflection angles (i.e., reasonably near line-of-sight) the diffraction field is not critically dependent on the analytical form of the surface provided it is sufficiently smooth. It can be expected that this will also be true for lossy obstacles.

8. Appendixes

8.1. Definition of Airy Integrals

The Airy integrals used in the above analysis can be represented by the following contour integrals

$$w_1(t)=\frac{1}{\sqrt{\pi}}\int_{\infty e^{i2\pi/3}}^{\infty} e^{\alpha}ds=u(t)-iv(t) \tag{26}$$

$$w_1'(t)=\frac{1}{\sqrt{\pi}}\int_{\infty e^{i2\pi/3}}^{\infty} se^{\alpha}ds=u_1'(t)-iv'(t) \tag{27}$$

$$w_2(t)=\frac{1}{\sqrt{\pi}}\int_{\infty e^{-i2\pi/3}}^{\infty} e^{\alpha}ds=u(t)+iv(t) \tag{28}$$

$$w_2'(t)=\frac{1}{\sqrt{\pi}}\int_{\infty e^{-i2\pi/3}}^{\infty} se^{\alpha}ds=u'(t)+iv'(t) \tag{29}$$

where $\alpha=st-s^3/3$.

The functions $u(t)$, $v(t)$, $u'(t)$, and $v'(t)$ have been tabulated by Fock [20] for real values of t.

8.2. Comparison With Results for a Spherical Model

The previous discussion has been based on a cylindrical model. It is of interest to compare these results with those of Fock [21] who employed a spherical model. To facilitate this discussion, it is desirable to review briefly Fock's work as it pertains to this particular problem. The notation will be changed to be consistent with the rest of the present paper.

The source is represented as a vertical electric dipole located at $r=a+h_1$ with respect to a sphere at $r=a$. (r, $\bar{\theta}$, ϕ are the usual spherical coordinates with origin at the center of the sphere and the axis in the direction of the dipole.) The great circle distance between the source and the receiver is denoted by s which, of course, is equal to $a\theta$. The receiver is located at $r=a+h_2$. In Fock's theory the following parameter is consistently employed

$$M=(ka/2)^{1/3}, \quad x=M(s/a)=M\bar{\theta},$$

$$y_1=kh_1/M \text{ and } y_2=kh_2/M.$$

The fields for this geometry can be derived from a single scalar function U which has the following form

$$U=e^{-ikR}/R+U^s,$$

where R is the distance from the source to the observer. The first term is a spherical wave and represents the primary influence whereas the quantity U^s is the secondary influence resulting from the presence of the spherical body. The rigorous analytical expression for U^s is well known, being given first by G. N. Watson in a form of a very slowly converging series. The transformation of this series into a complex integral had also been accomplished by Watson who suggested that a new series could be evolved if the contour was deformed round a set of complex poles. This results in the now well-known residue series which has been studied extensively. Rather than deforming the contour Fock decided that the complex integral could be employed directly. He approximated the spherical Bessel functions by an Airy integral approximation. This is equivalent to Van der Pol's and Bremmer's "Hankel approximation" which involved Hankel functions of order one-third. This representation is valid when the argument and the order of the spherical Bessel functions are very large, but nearly the same magnitude. This means that the heights h_1 and h_2 are small compared to the radius

a which itself must be very large compared to the wavelength. Fock is then able to write U in the following form

$$U=\frac{e^{-iks}}{\sqrt{sa\sin\theta}}V$$

where V is an attenuation function which is expressed conveniently as the sum of two parts, such that $V=V_1+V_2$, where V_1 does not depend on the electrical properties of the sphere and V_2 is dependent on these properties. The explicit representation for these integrals are

$$V_1=\left(\frac{ix}{\pi}\right)^{1/2}\left[\frac{1}{2i}\int_{C_1}e^{-ixt}w_1(t-y_2)w_2(t-y_1)dt+\int_{C_2}e^{-ixt}w_1(t-y_2)v(t-y_1)dt\right], \tag{30}$$

and

$$V_2=-\left(\frac{ix}{\pi}\right)^{1/2}\left[\frac{1}{2i}\int_{C_1}e^{-ixt}\frac{w_2'(t)-qw_2(t)}{w_1'(t)-qw_1(t)}w_1(t-y_1)w_1(t-y_2)dt\right.$$

$$\left.+\int_{C_2}e^{-ixt}\frac{v'(t)-qv(t)}{w_1'(t)-qw_1(t)}w_1(t-y_1)w_1(t-y_2)dt\right], \tag{31}$$

where $w_1(t)$, $w_2(t)$, and $v(t)$ are Airy integrals defined in appendix 8.1 and q is as defined previously. The contour C_1 is the straight line segment in the t plane from $\infty\exp(-i2\pi/3)$ to 0, and C_2 is the segment along the real axis from 0 to ∞. By some ingenious manipulations, the above integral representation V_1 is shown by Fock to correspond to Kirchhoff diffraction and consequently, V_2 can be regarded as the correction to the Kirchhoff theory. More explicitly

$$V_1\cong\frac{\sqrt{x}}{(y_1y_2)^{1/4}}e^{-i\omega_0}uF_0(\alpha)\text{ for }\alpha>0 \tag{32}$$

and

$$V_1\cong e^{-i\omega_0}\left[e^{i\alpha^2}-\frac{\sqrt{x}}{(y_1y_2)^{1/4}}uF_0(-\alpha)\right]\qquad\text{for }\alpha<0 \tag{33}$$

where $\omega_0\cong k(R-s)$, $u=(y_1y_2)^{1/4}/(\sqrt{y_1}+\sqrt{y_2})^{1/2}$ and $\alpha=u(x-\sqrt{y_1}-\sqrt{y_2})$. The function involves the Fresnel integral which is defined by

$$F_0(\alpha)=\frac{ei^{\left(\alpha^2+\frac{\pi}{4}\right)}}{\sqrt{\pi}}\int_\alpha^\infty e^{-ix^2}dx. \tag{34}$$

It is now possible to express V_2 in terms of the integral $G(X)$ encountered in the previous section. In fact,

$$V_2\cong-\frac{\sqrt{x}}{(y_1y_2)^{1/4}}e^{-i\omega_0}G(X), \tag{35}$$

where $X=\alpha/u=x-\sqrt{y_1}-\sqrt{y_2}$. Equations (33) and (35) given above for V_1 and V_2, respectively, are valid if x, y_1, and y_2 are all large compared to unity, but in such a way that X or $x-\sqrt{y_1}-\sqrt{y_2}$ is finite. It is also important to remember that these particular results are only valid for h_1 and h_2 small compared to a. This fact appears to have been overlooked in a recent paper by Kalinin [22]. In many physical situations one is interested in computing diffraction by a spherical surface when h_1 and h_2 are comparable or even large compared to a. It is quite easy, however, to relax this restriction if the proper asymptotic form of the spherical Hankel functions $h_{\nu-1/2}(\rho)$ is employed and ρ is either $a+h_1$, or $a+h_2$ in Fock's work. First it should be noted that

$$h_{\nu-1/2}(\rho)=\sqrt{\frac{\pi kb}{2}}H_\nu^{(2)}(\rho), \tag{36}$$

where $H_\nu^{(2)}(\rho)$ is the Hankel function of the second kind of order ν and argument ρ. The Debye second order representation for the Hankel function $H_\nu^{(2)}(kb)$ is now employed and leads to the approximation

$$h_{\nu-1/2}(\rho) \cong \frac{e^{-i(\xi-\pi/4)}}{(1-\nu^2/\rho^2)^{1/4}}, \tag{37}$$

where $\xi=(\rho^2-\nu^2)^{1/2}-\nu \cos^{-1}(\nu/\rho)$, being valid when $(\rho^2-\nu^2)>>\rho^{3/2}$. With this change, it is now possible to write the solution in the form

$$U=\frac{e^{-ikR}}{R} F, \tag{38}$$

where

$$F \cong [F_0(\alpha)+F_1]e^{-i\alpha^2}, \tag{39}$$

$$F_1=-G(X)/u, \tag{40}$$

$$u=\frac{a}{X}=\left(\frac{2ks_1s_2}{s_1+s_2}\right)^{1/2} \frac{(2/ka)^{1/3}}{2}, \tag{41}$$

and

$$X=(ka/2)^{1/3}\theta. \tag{42}$$

θ is the angular distance as usually defined, [8] and s_1 and s_2 are the linear distances measured from the source and the observer to the intersection of their tangent planes as in the previous section.

The research reported herein was sponsored in part by the U.S. Air Force Cambridge Research Center, Air Research and Development Command.

9. References

[1] J. C. Schelleng, C. R. Burrows and E. B. Ferrell, Ultra shortwave propagation, Proc. IRE **21,** 427 (March 1933).

[2] H. Selvidge, Diffraction measurements at ultra-high frequencies, Proc. IRE **29,** 10 (January 1941).

[3] K. A. Norton, M. Schulkin, and R. S. Kirby, Ground wave propagation over irregular terrain at frequencies above 50 Mc, Reference C to the report of the Ad Hoc Committee of the Federal Communications Commission for the evaluation of the radio propagation factors concerning the television and frequency modulation broadcasting services in the frequency range between 50 and 250 Mc (June 6, 1949).

[4] S. Matsuo, The method of calculating vhf field intensity and research on its variations, Report of the Electrical Communications Laboratory 621.30.001.6(047.3), Ministry of Telecommunications, Tokyo, Japan, (August 1950).

[5] F. H. Dickson, J. J. Egli, J. W. Herbstreit, and G. S. Wickizer, Large reductions of vhf transmission loss and fading by the presence of a mountain obstacle in beyond line-of-sight paths, Proc. IRE **41,** 967, (1953). See comments on this paper by J. H. Crysdale Proc. IRE **43,** 627 (1955).

[6] T. Kono, Y. Uesugi, M. Hirai, S. Niwa, and H. Irie, Measurement of field intensity of vhf radio waves behind Mt. Fugi, J. Radio Research Labs. (Tokyo) **1,** 1 (1954), Ministry of Postal Services, Tokyo, Japan.

[7] K. Bullington, Radio propagation at frequencies above 30 megacycles, Proc. IRE **35,** 1122 (1947).

[8] R. S. Kirby, H. T. Dougherty, and P. L. McQuate, Obstacle gain measurements over Pike's Peak at 60 to 1,046 Mc, Proc. IRE **43,** 1467 (1955).

[9] K. Furutsu, On the multiple diffraction by spherical mountains, J. Radio Research Labs. (Tokyo) **3,** 331 (1956).

[10] J. H. Crysdale, J. W. B. Day, W. S. Cook, M. E. Psutka, and P. E. Robillard, An experimental investigation of the diffraction of electromagnetic waves by a dominating ridge, IRE Trans. **AP–5,** 203 (1957).

[11] S. O. Rice, Diffraction of plane radio waves by a parabolic cylinder, Bell System Tech. J. **33,** 417 (1954).

[12] K. Artmann, Beugung polarisierten lichtes an blender endlicher dicke in gebiet der schattengrenze, Z. Physik **127,** 468 (1950).

---

[8] See figure 1c, for example. θ should not be confused with the spherical angle coordinate $\bar{\theta}$.

[13] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** 1480 (1955).

[14] H. E. J. Neugebauer and M. P. Bachynski, Diffraction by smooth cylindrical mountains, Proc. IRE **46,** 1619 (1958).

[15] I. P. Shkarofsky, H. E. J. Neugebauer, and M. P. Bachynski, IRE Trans. **AP–6,** 341 (1958).

[16] J. R. Wait and A. M. Conda, Pattern of an antenna on curved lossy surface, IRE Trans. **AP–6,** 348 (1958).

[17] H. Bremmer, Terrestrial radio waves, (Elsevier Publishing Co., Amsterdam, 1949). A comprehensive review of the early work of Watson, Van der Pol, and Bremmer is given.

[18] V. A. Fock, The field of a plane wave near the surface of a conducting body, J. Phys. U.S.S.R. **10,** 399 (1946).

[19] A. S. Goriainov, Diffraction of plane electromagnetic waves on a conducting cylinder, Radiotekh. i Elektron. **3,** 603 (1958). The paper, restricted to plane-wave incidence and perfect conductivity, is a limiting case of the present analysis.

[20] V. A. Fock, Diffraction of radio waves around the earth's surface (Academy of Sciences, U.S.S.R., 1946).

[21] V. A. Fock, Fresnel diffraction from convex bodies, Uspekhi Fiz. Nauk **43,** 587 (1951). A sphere model is used in a study restricted to antenna heights small compared to the radius of curvature.

[22] A. I. Kalinin, Approximate methods of computing field strength of ultra short waves, Radiotekhnika **12,** (1957). The quantity μ as used in his eq (28) is only valid if the antenna heights, h_1 and h_2, above the equivalent sphere are small compared to the radius. Clearly this is violated in Kalinin's application.

[23] K. Furutsu, Field strength in the vicinity of line of sight in diffraction, J. Radio Research Labs. (Tokyo) **3,** 55 (1956). (His functions $\chi_1^{(3)}$, $\chi_2^{(3)}$ etc. which are related to Airy integrals are usually approximated by the first term of their asymptotic expansions. It is difficult to see how this is justified near line of sight since the important values of z are not large in the "correction" terms. For example, in his eq (3.6) the leading term would vary as $\rho^{3/2}$ rather than $\rho^{1/2}$, if higher order terms of $\chi_1(z)$ and $\chi_2(z)$ are retained).

[24] T. T. Wu, The electromagnetic theory of light, Pt. I, Sci. Rept. No. 9, and T. T. Wu and S. R. Seshadri, Pt. II, Sci. Rept. No. 22, Cruft Lab., (Harvard Univ., 1958).

[25] J. B. Keller, Diffraction by a convex cylinder, IRE Trans. **AP–4,** 312 (1956).

[26] N. A. Logan, Fresnel diffraction by convex surfaces. Paper presented at Washington meeting of International Scientific Radio Union, May 4–7, 1959 (extensive numerical results for the care $q=0$ and ∞ and certain generalizations of the theory are reported in this paper).

BOULDER, COLO. (Paper 63D2–17).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Very-Low-Frequency Radiation Spectra of Lightning Discharges

W. L. Taylor and A. G. Jean

(April 9, 1959)

Spectral analyses are given of the groundwave portion of 33 sferic waveforms recorded from cloud-to-ground lightning discharges which occurred at distances ranging between about 150 and 600 kilometers from Boulder, Colo. Frequencies of peak energy lie between 5 and 20 kilocycles per second, which agree favorably with other published results. The average value of energy calculated from the groundwave pulses was found to be 26,600 joules, which is lower than values derived from other experiments. Various parameters, such as the peak amplitude and duration of the first half-cycle, are related to the radiated energy of the stroke.

1. Introduction

Appleton et al. (1926) were probably the first workers to utilize the cathode-ray oscilloscope in studying the electromagnetic energy emitted from lightning discharges. Following the development of radio communication techniques, it became evident that the lightning discharge was the source of interference to radio communication circuits.

In recent years, the radio signals emitted from lightning discharges, called atmospherics, or sferics, have been utilized as a source of signals in propagation studies at very-low radiofrequencies.

Various workers in England have been particularly active in the utilization of sferics in propagation studies, and it would be difficult and exhaustive to properly credit all of their accomplishments. However, as the work continued, it was learned that some atmospheric waveforms could be interpreted in terms of the reception of a series of pulses due to successive reflections between the earth and the ionosphere. It became evident that the character of the atmospheric waveform observed at great distances was materially altered by the propagation effects, while at shorter ranges, the waveform was more representative of the source function.

At the present time the use of atmospherics in propagation studies at the very-low radiofrequencies is common. NBS Boulder Laboratories have established a network of stations [1][1] for this purpose, and the material presented in this paper resulted from the simultaneous observation of atmospherics at the Bureau recording sites.

The objective of this paper is to present data on Fourier spectra of the electric field strength of signals radiated from return-stroke lightning discharges. The radiation spectra of lightning discharges were determined from observations of the groundwave portion of the atmospheric. The precautions taken in identifying and utilizing groundwave pulses for this purpose are described in this paper.

[1] Figures in brackets indicate the literature references at the end of this paper.

2. Equipment and Collection of Data

Stations at Boulder, Colo.; Salt Lake City, Utah; and Palo Alto, Calif.; were equipped to record the directions of arrival of atmospheric pulses, the pulse waveforms, and timing marks from which the times of arrival of individual pulses could be determined. Atmospheric waveforms used in the spectral analyses were recorded from vertical antennas. The overall amplitude response of the vertical antenna channel was constant within ± 1 db over the bandpass and sloped to 3-db cutoff points at frequencies of 1 and 100 kc. The phase response of this channel closely approximated a linear function of frequency within the bandpass.

A pair of vertical electrostatically shielded loop antennas, arranged at right angles to each other, were used in a direction-finding system at each station. The overall bandpass of each loop-antenna channel extended from 1 to 100 kc, and closely approximated that of the vertical-antenna channel. The direction-finding indications were obtained using the entire bandwidth of the loop-antenna channels. It was determined in earlier experiments [2] with wideband direction finders, that in many cases the groundwave component of the atmospheric could be discerned from the skywave components by virtue of differences in polarization and times of arrival of the groundwave and skywave pulses.[2] A time resolution of approximately 10 μsec is afforded by the wideband direction-finder, which is adequate to resolve the groundwave and skywave pulses. The direction of arrival of the atmospheric, as indicated by the vertically polarized groundwave component, should be virtually free from error due to the presence of horizontally polarized components in the atmospheric.

Timing marks derived from a secondary frequency standard, synchronized with the standard timing emissions from station WWV, were recorded on the atmospheric waveform records. Using these timing

[2] Excellent discussions regarding the location of sferic sources through the use of radio direction-finding have appeared in the literature; Horner [3], in particular, has discussed the probable sources of error in direction-finding, including that resulting from the reception of horizontally polarized components of ionospheric waves.

marks, the time of arrival of atmospherics could be determined at each station to an accuracy of about ±1 msec. Atmospheric waveforms resulting from a particular lightning discharge were located on the photographic records made at each station by virtue of their times of reception and directions of arrival.

The sweep of the oscilloscopes used in recording the atmospherics was activated by signals having amplitudes of about 50 mv/m or greater. Signals of smaller amplitude did not activate the sweep and therefore were not recorded. This feature provided some discrimination against atmospherics of relatively low amplitude; such as those of distant origin as well as some local cloud-to-cloud and precursory discharges. To minimize loss of early detail in the waveform, a 24-μsec delay line was used to delay recording the signal until after the sweep was initiated. The dynamic range of the recording equipment, in terms of the strength of a vertically polarized wave incident at the monopole antenna, extended from 0.2 to 6 v/m. Most of the analyzed atmospherics had peak field amplitudes approaching 1 v/m.

The locations of lightning discharges were determined by triangulation, using the direction of arrival indicated at each station. Due to the geometric arrangement of the 3-station network, the locations of lightning discharges which occurred near a recording station are relatively inaccurate; however, for distances greater than about 150 km, the error is estimated to be equal or less than 10 percent of the range.

It was desired to identify waveforms on the photographic records that resulted from vertical lightning discharges. Two tests that were used in an attempt to achieve this selection are described below. The relative amplitude of the vertical component of the electric field which would be registered at the appropriate ranges from a vertical discharge was calculated using range information available from the direction-finding "fixes," and from groundwave propagation curves in the literature [15]. The source was considered to be vertically polarized if the relative values of the observed and calculated pulse amplitudes agreed within about 20 percent.

If a vertically polarized wave arrives at grazing incidence, the direction-finder response will approximate a straight line inclined at an angle proportional to the azimuthal direction of arrival of the wave. An incident wave consisting of both horizontal and vertical components of the electric field will cause an elliptical direction-finding pattern to be produced. The direction-finder responses to the atmospheric pulses which were selected for analyses indicated that the groundwave pulses were vertically polarized, and the first-hop skywave pulses were essentially, but not entirely, vertically polarized. Atmospheric waveforms which satisfied both the "amplitude" and "ellipticity" criteria outlined above were assumed to result from vertical lightning discharges.

No sferics were analyzed which originated at distances less than 150 km from Boulder, due, in part, to limitations in the accuracy of "fixes" at short ranges, and also to minimize contributions by the induction field at frequencies near 1 kc. It was desirable to limit the maximum observation range in order to minimize the distortion of the groundwave pulse by propagation and to reduce the interference caused in the groundwave by the first-hop skywave pulse.

The waveforms of two sferics, representative of those used in this analysis, are reproduced in figure 1a. The range to the source is approximately 165 km from both sferics 1 and 2. The groundwave portion of sferic 1 is believed to comprise the first 110 μsec of the approximate 450-μsec time base. The earliest evidence of the arrival of the first-hop skywave pulse in sferic 1 occurs at 147 μsec. Note that the initial half cycle of the skywave pulse is negative, or opposite to the polarity of the first half-cycle of the groundwave pulse. This reversal in polarity of the initial part of the first-hop skywave pulses evident in figures 1a and 1b is a consequence of the pseudo Brewster angle in ionospheric reflection. This phenomenon was discussed in an earlier paper [1].

The height of the ionospheric reflection was calculated to be about 65 km, corresponding to a range of 165 km and a 147-μsec time interval between the reception of the groundwave and the first-hop skywave pulses. In this calculation, the ionosphere was assumed to be a homogeneous ionized medium with a sharp lower boundary [4].

Note that sferic 2 is similar in character to sferic 1, but the groundwave is of larger amplitude and longer duration. The first-hop skywave pulse was received before the entire groundwave pulse was recorded. It was necessary in such cases to complete the groundwave pulse as indicated by the dotted line before undertaking the spectral calculations. Since a very small fraction of the total area enclosed by the waveform is contained under the sketched-in portion, the error introduced into the spectral calculations is small.

Waveforms of sferics 17 and 18,[3] which were received at ranges of 460 km under nighttime conditions, are given in figure 1b. Note that in these

[3] Numbers were assigned to the atmospherics in order of their selection for analyses and each number associates an atmospheric with its spectrum. Sferics 1, 2, 17, and 18 were selected to represent the variability encountered in the reception of the groundwave pulses. The waveforms of the remaining atmospherics (sferics 3 through 16 and 19 through 33) are not presented in this report.

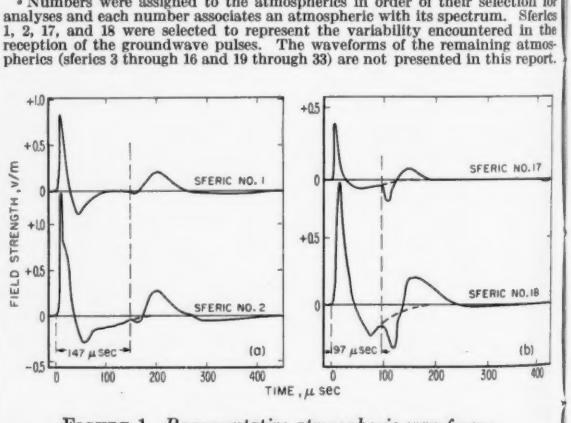


FIGURE 1. *Representative atmospheric waveforms.*

cases the time interval between reception of the groundwave and the first-hop skywave pulses is about 97 μsec compared to an interval of 147 μsec in figure 1a. As a consequence of the shorter time interval between the reception of the groundwave and first-hop skywave pulses, a larger portion of the groundwave pulses are concealed by the skywave pulses than for the sferics 1 and 2. Prior to performing the spectral analyses of the groundwave pulses for sferics 17 and 18, the trailing edges of the groundwave pulses were sketched in as indicated. A larger error resulted in such cases; e.g., the area under the sketched-in portion in sferic 18 is approximately 13 percent of the total area enclosed by the waveform of the groundwave pulse; however, this error is not considered to be serious in the present analyses.

The ionospheric reflection height for these nighttime observations was calculated to be about 80 km, using a range of 460 km and a time interval between reception of groundwave and first-hop skywave pulses of 97 μsec.

3. Data Analyses

Let $G(t)$ represent the instantaneous value of the vertical component of the electric field in units of volts per meter. Also let the waveform start at time $t=0$, so that

$$G(t)=0, \quad \text{for } t\leq 0,$$

and the end of the pulse occurs at $t=\tau$ so that

$$G(t)=0, \quad \text{for } t\geq \tau.$$

The complex spectrum of the pulse can be represented by the Fourier integral

$$F(f)=\int_0^\tau e^{-j2\pi ft}G(t)dt, \tag{1}$$

where f represents the frequency in cycles per second and t is the time in seconds after the beginning of the pulse.

The complex spectral components of the atmospheric were calculated at intervals of 1 kc on an electronic computer. The absolute value, $|F(f)|$, of the spectra of the groundwave pulses of individual atmospherics are given in figures 2a through 2h. Fourteen of the analyzed pulses were received under daylight conditions from a storm approximately 165 km from Boulder, and 19 pulses were received under nighttime conditions from ranges varying between 400 and 620 km from Boulder. The approximate range at which each pulse was observed is indicated on the appropriate figure.

The frequencies of maximum energy for return stroke lightning discharges have been reported by a number of different workers.[4] Their separate findings indicate that the frequency peaks occur in the 5- to 20-kc region. It will be observed that the peak frequencies in figure 2 fall within this range of frequencies.

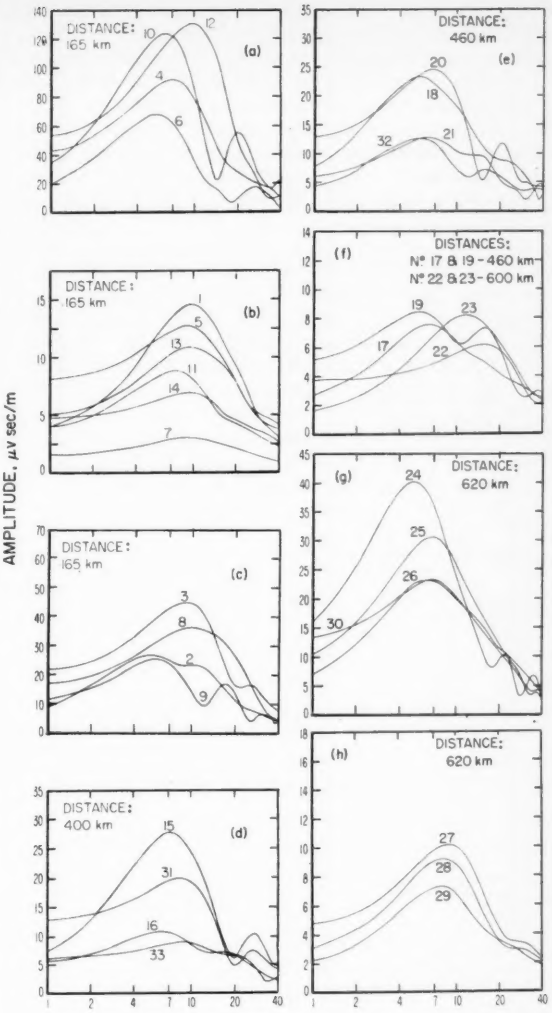


FIGURE 2. *Amplitude spectra of atmospherics (see text).*

With a knowledge of the complex spectrum $F(f)$ of a groundwave pulse observed at distance d_1, and of the appropriate values of earth constants, it is possible to calculate the spectrum $|F(f)|$ of a groundwave pulse at a different range, d_2, using methods evolved by Wait [14]. This rigorous method was not followed in this paper, due to the complex nature of the calculations and the lack of adequate information concerning earth constants. The less rigorous method used in normalizing the spectra $|F(f)|$ to a common distance is described below.

The atmospheric pulses were recorded at distances between 150 and 600 km and their spectral peaks

[4] It is difficult to adequately reference all of the available literature dealing with this subject. A few representative references [5–12] have been selected.

were attained at frequencies near 10 kc. A uniform earth, having a conductivity of about 10^{-2} mho/m, was assumed to exist along the propagation paths. Referring to the curves of groundwave amplitude versus distance [15], it is evident for these values of earth conductivity, distance, and frequency, that the signal amplitude decreases approximately at the inverse distance rate. Accordingly, the individual spectra were referred to a common distance of 1 km by increasing each spectral component in proportion to the observation range. The spectrum shown in figure 3 is the average of the individual spectra normalized to this distance. The deviation of the individual spectra from the average is not very large which seems to support evidence in the literature [12, 13] that the radiation spectra of return-stroke lightning discharges do not vary a large amount from event to event. The 1-km value was selected as a matter of convenience; the spectrum in figure 3 is not intended to represent the actual distribution of amplitude with frequency at that distance.

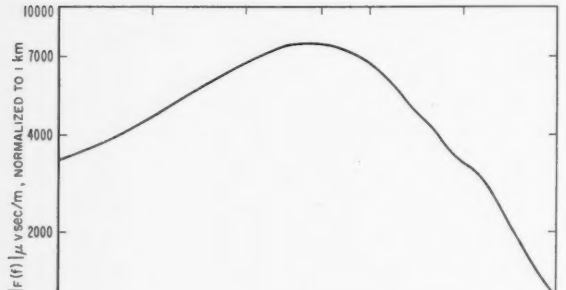


FIGURE 3. *Average of amplitude spectra, normalized to 1 km.*

The peak amplitude of this spectrum occurs at a frequency of 6.5 kc; at frequencies below the peak, the spectral amplitude is proportional to the 0.4 power of the frequency, while at frequencies higher than the peak, the spectral amplitude decreases as the inverse 1.5 power of the frequency.

The total energy available in a lightning discharge has been estimated to be about 2×10^{17} ergs [16]. This large amount of energy is partitioned into a number of different forms; for example, some energy is required to ionize the gas in the discharge channel, a larger portion is liberated as frictional heat in the movement of gaseous components, and some is evident as electromagnetic radiation, etc. An appreciable part of the energy is consumed in "ground losses" due to the flow of induced currents in the earth [17].

For the purposes at hand, it was desired to estimate the amount of vlf radio energy radiated into the earth's atmosphere (i.e., exclusive of earth losses) by the lightning discharges. This form of energy is a very small portion of the total amount available.

The energy density, E_d, of the transient $G(t)$ is represented in joules/meter^2 by

$$E_d=\frac{1}{120\pi}\int_0^\infty G(t)^2dt. \tag{2}$$

It was assumed that the current at the source was a simple linear dipole oriented along the vertical axis. The radiation component of the dipole electric field is proportional to cos θ, where θ is the angle of elevation measured from the horizontal. The vlf energy, E_T, in joules, radiated into the atmosphere was obtained by integrating the energy density E_d over the surface of the radiation pattern of the source. Thus,

$$E_T=\int_0^{2\pi}\int_0^{\pi/2} E_d r^2\cos^3\theta d\theta d\phi \tag{3}$$
$$=4/3\pi r^2 E_d,$$

where θ is the angle of elevation measured from the horizontal, ϕ is the angle measured in the equatorial plane, and r is the distance in meters from the source to the point of observation.

Values of energy represented by each of the pulses, calculated using expression (3), are given in table 1.

TABLE 1. *Very-low-frequency energy radiated into the atmosphere by return-stroke lightning discharges*

Atmospheric No.	Distance	f_0	E_T
Sferic groundwave recorded at Boulder during daylight July 7, 1959			
	km	*kc*	*j*
1	165	10.0	2.2×10^{3}
2	165	5.0	6.2×10^{3}
3	165	9.0	1.7×10^{4}
4	165	7.0	5.5×10^{4}
5	165	9.0	1.8×10^{3}
6	165	5.5	2.1×10^{4}
7	165	9.5	1.1×10^{2}
8	165	10.0	1.5×10^{4}
9	165	6.0	4.3×10^{3}
10	165	6.0	8.8×10^{4}
11	165	7.5	6.9×10^{2}
12	165	9.5	1.3×10^{5}
13	165	10.0	1.5×10^{3}
14	165	9.5	5.9×10^{2}
Sferic groundwave recorded at Boulder during nighttime July 26, 1958			
15	400	7.5	2.1×10^{4}
16	400	6.0	6.5×10^{3}
31	400	8.5	2.0×10^{4}
33	400	8.5	5.5×10^{3}
17	460	6.5	4.5×10^{3}
18	460	5.5	2.9×10^{4}
19	460	5.5	6.2×10^{3}
20	460	7.0	3.1×10^{4}
21	460	7.0	1.1×10^{4}
32	460	5.5	9.3×10^{3}
22	600	15.5	7.9×10^{3}
23	600	12.0	1.1×10^{4}
24	620	5.5	1.1×10^{5}
25	620	7.0	9.2×10^{4}
26	620	6.5	6.4×10^{4}
27	620	9.0	1.3×10^{4}
28	620	8.5	9.6×10^{3}
29	620	8.0	6.4×10^{3}
30	620	6.5	6.3×10^{4}

The average of these values of energy is 26,600 j. On another occasion [10], sferics were received from thunderstorms in a frontal region extending through Texas, Louisiana, and Oklahoma. The average of the total radiated energy values calculated from seven of these sferics was about 250,000 j; the individual values range from 130,000 to 500,000 j. Laby, et al. [11], reported the average value of energy radiated from 35 sferics observed in Australia was 200,000 j. It appears that the sferics used in the present work resulted from less energetic discharges than those reported by Taylor and by Laby. It is not at all certain that the energy value reported here is representative of return-stroke discharges from thunderstorms occurring in the Rocky Mountain area. However, part of the differences between the energy values quoted might have resulted from differences in thunderstorm characteristics, differences in topography, earth conductivity, etc.

The importance of land elevation in determining thunderstorm characteristics was pointed out by Loeb [18]. In general, the lightning discharges used in this analysis occurred over land of higher elevation than those previously reported by Taylor. It is not known in what area of Australia the lightning flashes occurred that were used in Laby's analysis.

It is interesting to note that the ratio of the total energy available in a lighting discharge to the values of electromagnetic energy radiated into the atmosphere is as large as 10^5 or 10^6.

A wide variation was noted in the detail of the individual groundwave pulses; however, most of them consisted of two half-cycles. The first half-cycles were larger in amplitude and shorter in duration than the corresponding values in the second half-cycle. These two characteristics of the first half-cycle would be expected to control the total radiated energy to a certain extent. After normalizing the observations to a 1-km range, the relationships existing between various parameters mentioned above were examined. Some of these comparisons are described below.

(A) *Frequency of spectral peak versus total radiated energy.* The relationship between the frequency of spectral peak (f_0) and the total radiated energy (E_T) is given in figure 4. Although the data points scatter, there is a slight tendency for the more energetic sferics to attain energy peaks at the lower frequencies.

(B) *Peak amplitude of sferics versus total radiated energy.* The peak amplitude of the groundwave pulse (E_{max}) is plotted as a function of total radiated energy in figure 5. There is a close relationship between these parameters. The straight line drawn through the data points relate total radiated energy to the 2.4 power of the peak field strength.

(C) *Spectral peak versus total radiated energy.* The relationship between the peak of the spectrum of the groundwave pulse, $F(f)_{max}$, versus the total radiated energy of the pulse is given in figure 6. A close relationship is indicated between these parameters. The straight line through the data points indicates that the energy is proportional to the 1.8 power of the spectral peak. The frequency of the spectral peak is not explicitly accounted for in this relationship.

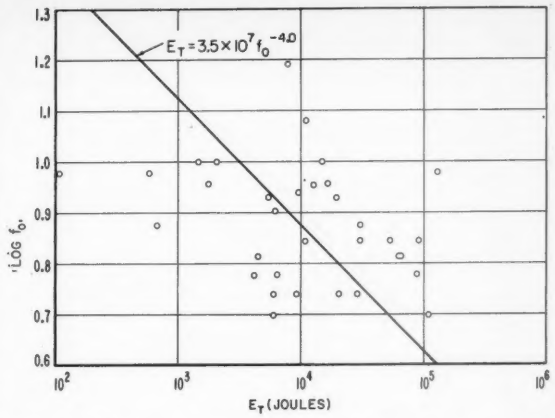


FIGURE 4. *Frequency of spectral peak versus total radiated energy.*

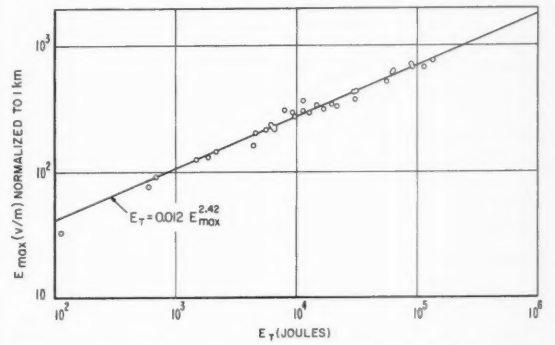


FIGURE 5. *Pulse amplitude versus total radiated energy.*

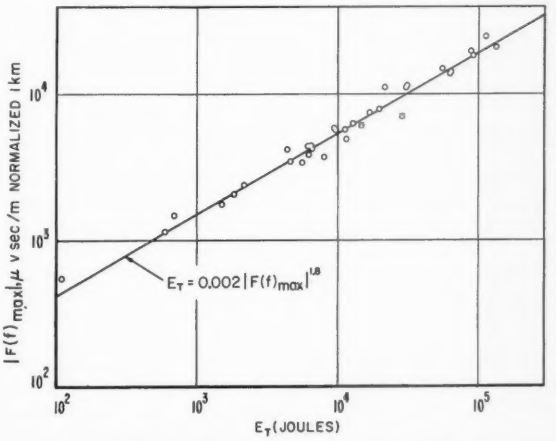


FIGURE 6. *Spectral peak versus total radiated energy.*

(D) *Length of first half-cycle versus total radiated energy.* In figure 7 the relationship between the length of the first half-cycle of the groundwave pulse

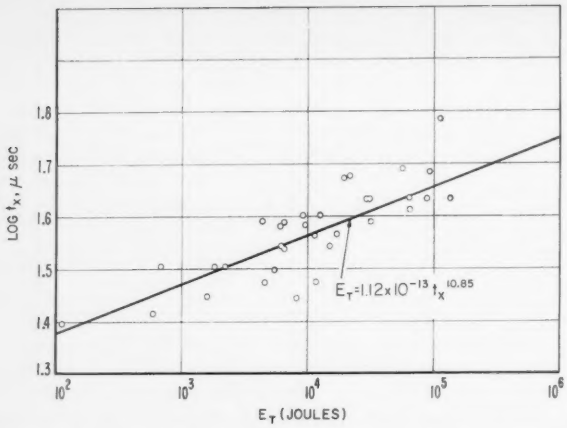


FIGURE 7. *Length of first half-cycle versus total radiated energy.*

(t_x in microseconds) and the corresponding value of total radiated energy is plotted. The relative importance of the first half-cycle in determining the energy is indicated in this plot and in figure 5, in which the peak amplitude of the groundwave pulse is plotted as a function of total radiated energy.

The field strength of a pulse is related to the rate of change of the source moment [13, 14, 19], hence, flashes having rapidly rising currents would be expected to produce sferics of high peak amplitude. The duration of the first half-cycle of the groundwave pulse should correspond to the time at which the source current reached its maximum value. Hence, sferics of large energy result from rapidly rising and long continuing source currents. It would be expected that the more energetic discharges would produce sferics of longer duration, and a strong tendency of this nature is observed in the data; however, it was also observed that sferics representing relatively low values of energy attain spectral peaks within the same frequency region, 5–20 kc, as those of larger energy.

The close relationship existing in figure 5 between the peak amplitude of the first half-cycle and total radiated energy suggests the possibility of gaining a useful estimate of energy by observing the peak amplitude of the groundwave pulse and by determining the distance to the flash. However, it is not known that this relationship will hold equally well at the higher energy levels.

It would be of interest to compare the characteristics of sferics originating (1) over land of different elevations and those originating over sea water, and (2) from frontal and air-mass thunderstorms. This would aid in appraising the influence of topography, earth conductivity, and meteorological factors in determining the characteristics of sferics.

Atmospheric waveforms recorded at Boulder were used in the present analysis. Simultaneous recordings of these atmospherics are also available from the other stations located at Stanford and Salt Lake City. Analyses of this data is continuing to determine values of propagation attenuation [10] and phase velocity of propagation [20, 21].

The authors wish to acknowledge the mathematical calculations made by L. Jerome Lange and A. Murphy, the preparation of the programing for the computer by H. H. Howe, and helpful conversations with Roger Gallet and James R. Wait during the preparation of this paper.

4. References

[1] A. G. Jean, L. J. Lange, and J. R. Wait, Ionospheric reflection coefficients at vlf from sferic measurements, Geofisica Pura e Applicata **38,** 147 (1957).
[2] R. A. Helliwell, A. G. Jean, and W. L. Taylor, Some properties of lightning impulses which produce whistlers, Proc. IRE **46,** 1760 (1958).
[3] F. Horner, The accuracy of the location of sources of atmospherics by radio direction finding, Proc. Inst. Elec. Engrs. (London) **101, Pt. III,** 383 (1954).
[4] J. R. Wait and A. Murphy, Multiple reflections between the earth and the ionosphere in vlf propagation, Geofisica Pura e Applicata **35,** 61 (1956).
[5] C. E. R. Bruce and R. H. Golde, The lightning discharge, J. Inst. Elec. Engrs. (London) **88, Pt. II,** 487 (1941).
[6] H. Norinder, Variations of the electric field in the vicinity of lightning discharges, Arkiv Geofysik **1,** 543 (1953).
[7] P. W. A. Bowe, The waveforms of atmospherics and the propagation of very low frequency radio waves, Phil. Mag. [7] **42,** 121 (1951).
[8] F. W. Chapman and A. G. Edwards, Some observations on the frequency spectrum of atmospheric disturbances due to thundercloud discharges and their effect on resonant circuits, Proc. Union Radio Scientifique Internationale Comm. IV, No. 252 (1950) Zurich.
[9] A. Kimpara, The typhoon kezia and atmospherics, Proc. Japan Acad. **27,** 366 (1951).
[10] W. L. Taylor and L. J. Lange, Some characteristics of vlf propagation using atmospheric waveforms, Second Conf. Atmospheric Elec. May 1958 (to be published April 1959).
[11] T. H. Laby, J. J. McNeill, F. G. Nicholls, and A. F. B. Nickson, Waveform energy and reflexion by the ionosphere of atmospherics, Proc. Roy. Soc. (London) [A] **174,** 145 (1940).
[12] F. Chapman and R. C. V. Macario, Propagation of audio-frequency radio waves to great distances, Nature **177,** 930 (1956).
[13] A. D. Watt and E. L. Maxwell, Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (1957).
[14] J. R. Wait, On the waveform of a radio atmospheric at short ranges, Proc. IRE **44,** 1052 (1956).
[15] J. R. Wait and H. H. Howe, Amplitude and phase curves for ground-wave propagation in the band 200 cycles per second to 500 kilocycles, NBS Circ. 574 (May 1956).
[16] C. W. Allen, Astrophysical quantities, p. 123 (Univ. of London, Athlone Press, 1955).
[17] J. R. Wait, personal communication.
[18] L. B. Loeb, Thunderstorms and lightning strokes, p. 330 Modern physics for the engineer by Royal Weller and others (McGraw-Hill Book Co., New York, 1954).
[19] H. A. Thomas and R. E. Burgess, Survey of existing information and data on radio noise over the frequency range 1–30 Mc/s, DSIR Radio Research, Special Rept. 15, p. 38 (1947).
[20] J. R. Wait, Propagation of very-low-frequency pulses to great distances, J. Research NBS **61,** 187 (1958) RP2898.
[21] Ia. L. Al'Pert and S. V. Borodina, Zhur. Eksptl. i Teoret. Fiz. (J. Exp. Theoret. Phys. U.S.S.R.) **33,** 1305 (1957).

BOULDER, COLO. (Paper 63D2–18).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September-October 1959

Radio-Wave Scattering by Tropospheric Irregularities

Albert D. Wheelon [1]

(April 16, 1959)

The subject of radio-wave scattering by turbulent irregularities in the tropospheric index of refraction is reviewed. Descriptions of the turbulent medium are considered first, together with a status report on physical theories for the spectrum of irregularities. Phase and amplitude scintillations induced on electromagnetic waves propagated along line-of-sight paths are discussed next. The complementary problems of signal statistics and electromagnetic-propagation calculations are summarized and compared with available data. Both the geometrical optics and wave theory approaches are discussed. The theory of propagation beyond the horizon by scattering from the same irregularities is then described. A current review of the predictions for received power, signal fluctuations, and antenna effects is given. The paper is primarily a review of the essential and auxiliary predictions of scatter theory, but also contains a considerable amount of unpublished research by the author.

1. Introduction

The existence of random irregularities in the atmospheric index of refraction has been well established by direct measurements with microwave refractometers [1, 2].[2] The influence of these refractive irregularities on electromagnetic waves is confirmed by the random fading of microwave signals received beyond the optical horizon. Stochastic phase and amplitude variations imposed on electromagnetic waves propagated along line-of-sight paths confirm this cause-effect relationship. This is not to deny the influence of other mechanisms (layers, ducts, et al.) which may also contribute to scatter-field strengths. It *does* certify that the electromagnetically predicted relation between (turbulent) refractive variations and signal characteristics is indeed established.

The entire subject of electromagnetic propagation through turbulent refractive irregularities has received considerable theoretical and experimental attention during the post-war era. This attention is due to the enormous importance of "scatter propagation" communication systems on the one hand, and a basic limitation imposed on radio tracking and guidance systems for space vehicles on the other. The subject also has considerable scientific interest in its own right, particularly insofar as it unites electromagnetic theory, turbulence theory, and probability in an essential mixture. Once established, the influence of atmospheric irregularities on electromagnetic waves presents a valuable tool for studying turbulence at very large Reynolds numbers by simple radio means.

It now seems appropriate to try to establish how far the considerable volume of original research has brought the subject. This is particularly important since the basic concept is periodically brought under attacks of various sorts, and it is desirable to try to set forth the essential and auxiliary predictions of the basic theory in a coherent, unambiguous fashion. More important, however, is the need to present a balanced status report on the theoretical predictions, against which accumulating experimental results can be ranged to establish the true role of this mechanism in an objective manner.

The present paper is primarily a review of theoretical research on line-of-sight and scatter propagation in the United States, with a short status report on turbulence theories of atmospheric irregularities. It is not a genuine review, in that it leans rather heavily on the author's own research as a basis for unifying diverse results. Furthermore, a substantial amount of unpublished material is included herein. The report goes one step further and points out a number of outstanding problems which are solved in principle, but not in detail. A detailed, balanced comparison of these theoretical results with experimental data was not attempted for this broad field, although occasional references to significant results and prior comparisons are indicated by way of stimulating this crucial testing.

The substance of this report can be summarized briefly. In so far as the single scattering (Born) approximation and/or geometrical optics are valid representations of the electromagnetic response of propagating waves to refractive irregularities, the entire subject is solved in principle. The geometrical-propagation effects can be computed in any desired detail (or at least be reduced to integrals), independent of the turbulence model employed. This important separation is achieved by using the spectral representation for the space correlation of refractive irregularities, which allows any spectrum model to be combined with any propagation model. The more difficult electromagnetic problem of multiple scattering has not been considered adequately and is not discussed here.

[1] Consultant, Boulder Laboratories, NBS; permanent address, Space Technology Laboratories, Inc., Los Angeles 45, Calif.
[2] Figures in brackets indicate the literature references at the end of this paper.

2. Description of the Turbulent Medium

2.1. Statistical Characterization

The tropospheric index of refraction varies from point-to-point in a random manner. To describe this effect, one decomposes the dielectric constant ($\epsilon=n^2$) into its mean value and a small stochastic component.

$$\epsilon(\vec{r}, t)=\epsilon_0+\Delta\epsilon(\vec{r}, t). \tag{2.1.1}$$

It is this quantity which enters into the basic electromagnetic equations. ϵ_0 is usually a function of altitude (h) only, and its gradient appears in certain physical theories of turbulent irregularities. The stochastic function $\Delta\epsilon(\vec{r},t)$ is treated as a stationary, zero-mean random process in space and time, and is usually assumed to be spatially homogeneous and isotropic. If brackets denote ensemble, space, or time averages (which are equivalent for a statistically stationary, homogeneous atmosphere) the first two moments of $\Delta\epsilon$ are:

$$\langle\Delta\epsilon(\vec{r}, t)\rangle=0 \tag{2.1.2}$$

$$\langle\Delta\epsilon(\vec{r}, t)\Delta\epsilon(\overrightarrow{r+R}, t)\rangle=\langle\Delta\epsilon^2\rangle C(|\vec{R}|). \tag{2.1.3}$$

The intensity of dielectric fluctuations $\langle\Delta\epsilon^2\rangle$ is measured experimentally [1, 2] as several parts in 10^{-12}. The space-correlation function $C(R)$ falls to $1/e$ in a distance of several hundred feet, which is related to the scale length l_0 of the irregularities. Little attention has been paid to higher moments of $\Delta\epsilon$, except to note that they are completely specified in terms of $C(R)$ if the $\Delta\epsilon$ form a Gaussian random process.

2.2. Space Correlation Models

The space-correlation function $C(R)$ is the fundamental characterization of the turbulence. First-order propagation quantities depend upon integrals of $C(R)$, and one must assume an analytic model to perform the relevant calculations. The correlation must have the form

$$C(R)=f(R/l_0), \tag{2.2.1}$$

where $f(x)$ is the Fourier transform of a non-negative function. l_0 is called the scale length or average size of the blobs in the hierarchy of decaying eddies. The popular models for $f(x)$ are plotted together in figure 1.

1. *Bessel:* $f(x)=xK_1(x)$

This correlation model was devised to predict: (a) The observed linear-wavelength dependence of average power for tropospheric scatter propagation; and (b) certain features of line-of-sight propagation experiments. It has zero slope at the origin, and

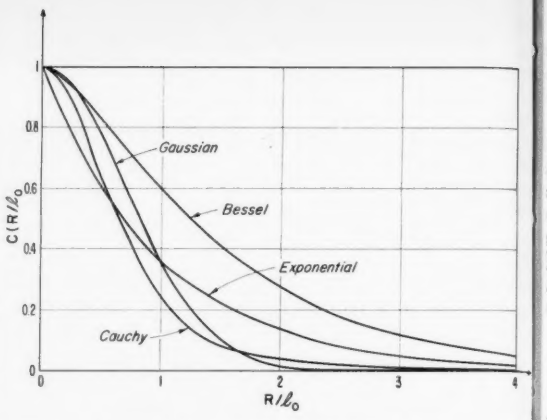


FIGURE 1.—*Plot of four common space correlation models for isotropic atmospheric irregularities.*

is suggested by the theory of turbulent mixing-in gradient.

2. *Exponential:* $f(x)=e^{-x}$

This function has been widely used to study scatte propagation. However, its cusplike behavior (finit slope) at the origin implies infinite values for angle-of arrival errors and other phase-derivative processes The exponential model is possibly suggested by th physical theory of turbulent pressure fluctuations.

3. *Gaussian:* $f(x)=e^{-x^2}$

This is widely used for line-of-sight calculation because it can be integrated easily. It does no predict the results of tropospheric scatter experiment adequately and apparently has no foundation i physical theories of turbulence.

4. *Cauchy:* $f(x)=[1+x^2]^{-2}$

This model has been used for line-of-sight ra theory calculations, but is not founded in physica theories of turbulent irregularities.

It should be emphasized that any correlation functio is only a model of the atmospheric irregularities, an must be carefully evaluated in terms of its ability t predict the results of experiments.

It is possible to generalize the foregoing to include anisotropic and inhomogeneous turbulent structures. Anisotropic irregularities can be described by introducing different scale lengths in orthogonal directions, viz,

$$C(R)=f\left(\left[\frac{x^2}{l_x^2}+\frac{y^2}{l_y^2}+\frac{z^2}{l_z^2}\right]^{1/2}\right), \tag{2.2.2}$$

and such effects are apparently relevant in the troposphere.

Refractometer experiments show that the intensity of dielectric fluctuations decreases with altitude [1]. One can describe an inhomogeneous field of this sort with a height-dependent $\langle\Delta\epsilon^2\rangle$ in eq (2.1.3), so long as the intensity does not change significantly in the several scale lengths l_0 required for the (space) averages.

A more difficult problem is encountered with the assumption of stationarity. There now seems to be experimental evidence [3] that the dielectric variations either: (1) do *not* form a stationary process; or (2) that there is a great deal of power in the low-frequency part of the spectrum of $\Delta\epsilon$, which appears as trend when viewed over a short span of time or space. Both explanations are probably partially correct, with a rapidly fading stationary process superimposed on long-period (diurnal, seasonal) harmonic components. Unfortunately, there is no adequate theory of nonstationary processes available to describe such effects.

2.3. Turbulence Spectrum Method

An equivalent characterization of a turbulent medium is given by the spectrum of irregularities $S(\vec{k})$ [4, 5], which is the Fourier decomposition of the space correlation function,

$$\langle \Delta\epsilon^2 \rangle C(\vec{R}) = \frac{1}{8\pi^3} \int d^3k \epsilon^{-i\vec{k}\cdot\vec{R}} S(\vec{k}). \qquad (2.3.1)$$

This representation has the advantge of exhibiting an explicit R-dependence for the correlation without committing the study to a particular turbulence model. Propagation calculations can be performed on the generic term $\exp[-i\vec{k}\cdot\vec{R}]$ by reversing the orders of integration. The physics of the atmospheric irregularities re-enters the problem eventually in the spectrum $S(\vec{k})$, which must be introduced to complete the wavenumber (k) integration. If the medium is isotropic, $S(\vec{k})$ depends only on $|\vec{k}|$, and the space correlation becomes:

$$\langle \Delta\epsilon^2 \rangle C(R) = \frac{1}{2\pi^2} \int_0^\infty dk k S(k) \frac{\sin kR}{R}. \qquad (2.3.2)$$

The mean square fluctuation at a point,

$$\langle \Delta\epsilon^2 \rangle = \frac{1}{2\pi^2} \int_0^\infty dk k^2 S(k), \qquad (2.3.3)$$

indicates that the spectrum represents the ability of each blob-size group (wavenumber range) to produce irregularities.

Equation (2.3.1) may be inverted to establish the spectrum which corresponds to each of the correlation functions discussed earlier. Table 1 indicates the spectra for the Bessel, Exponential, Gaussian, and Cauchy models. The small R behavior of the correlation is determined by the large k range of the spectrum and vice versa.

TABLE 1. *Correspondence of turbulence spectrum and space correlation functions for several models of atmospheric irregularities*

	Bessel	Exponential	Gaussian	Cauchy
Space correlation $C(R)$	$\frac{R}{l_0} K_1\left(\frac{R}{l_0}\right)$	ϵ^{-R/l_0}	ϵ^{-R^2/l_0^2}	$\left[1+\frac{R^2}{l_0^2}\right]^{-2}$
Turbulence spectrum $S(k)$	$6\pi^2 \frac{\langle \Delta\epsilon^2 \rangle l_0^3}{[1+k^2 l_0^2]^{5/2}}$	$8\pi \frac{\langle \Delta\epsilon^2 \rangle l_0^3}{[1+k^2 l_0^2]^2}$	$\pi^{3/2} \langle \Delta\epsilon^2 \rangle l_0^3 \epsilon^{\frac{-k^2 l_0^2}{4}}$	$\pi^2 \langle \Delta\epsilon^2 \rangle l_0^3 \epsilon^{-kl_0}$

2.4. Physical Theories of Dielectric Irregularities

The index of refraction in the (nonionized) troposphere and stratosphere is independent of radio-frequency from 10 to 10,000 Mc. Variations of the refractive index from unity are measured as several hundred parts per million, and are related to standard meteorological parameters by the empirical formula:

$$n-1 = \left\{ 77.6 \frac{P}{T} + 3.73\ 10^5 \frac{e}{T^2} \right\} 10^{-6} \qquad (2.4.1)$$

where

T = air temperature is degrees Kelvin (absolute),
P = total air pressure in millibars, and
e = partial water-vapor pressure in millibars.

Dielectric variations ($\epsilon = n^2$) are thus related to temperature, pressure and water-vapor fluctuations by

$$\Delta\epsilon = 2\ 10^{-6} \left\{ \begin{array}{l} -\Delta T \left(77.6 \frac{P}{T^2} + 7.46\ 10^5 \frac{e}{T^3} \right) \\ +\Delta P \left(\frac{77.6}{T} \right) + \Delta e \left(\frac{3.73\ 10^5}{T^2} \right) \end{array} \right\} \qquad (2.4.2)$$

Dielectric variations in the lower troposphere are governed principally by water-vapor irregularities, and one is dealing with the irregularities established in a passive scalar by turbulent movements of the random atmospheric velocity field. This process is characterized by the turbulent mixing theories described below.

In the stratosphere there is a negligible amount of water vapor, and dielectric variations are produced almost exclusively by thermodynamic fluctuations: ΔT and ΔP. The primary mechanism for producing irregularities in the stratosphere is also probably turbulent mixing by rotating eddies. However, one must now recognize thermodynamic expansion (perhaps adiabatic [6]) of the gas as the eddies transfer parcels of air between different altitude levels, corresponding to different ambient temperatures and pressures. This expansion correction is not dominant in the lower troposphere, but probably plays a central role in the propagation of scatter signals to distances greater than 700 km by common volume elements in the stratosphere [6].

The spectrum method makes direct contact with physical theories or turbulent irregularities which predict $S(k)$ in certain wavenumber ranges. These theories identify three distinct ranges of the spectrum, as shown in figure 2.

1. $0 < k < k_0$: The input or blob-creating range corresponds to very large scale lengths, which are probably anisotropic. The upper end, k_0, of this

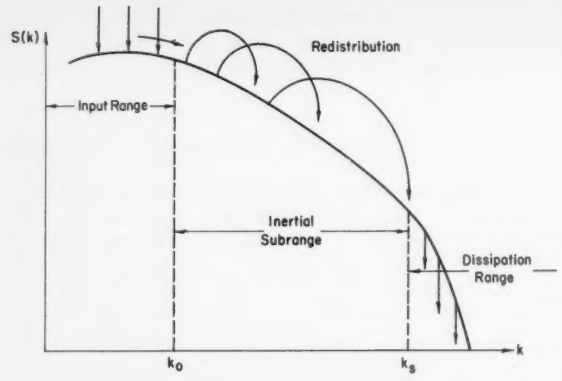


FIGURE 2.—*Typical spectrum of dielectric irregularities plotted versus wavenumber, showing various turbulence ranges.*

range is approximately the reciprocal of the correlation or scale length l_0 discussed earlier, which is of the order of 100 m in the troposphere. Since we do *not* know (in detail) how turbulence is created from a laminar flow, turbulence theories carefully avoid this range. It is important to note that the spectra listed in table 1 *do* make definite statements about this range, so the large R behavior of the corresponding correlation models must be regarded as extrapolation or pure guesswork.

2. $k_0<k<k_s$: The inertial subrange is characterized by redistribution of the turbulent "energy" toward higher wavenumbers and represents the progressive subdivision of eddies in the turbulent structure. Dimensional analysis can be used here with considerable confidence. Tropospheric-scatter propagation depends principally on this inertial range.

3. $k_s<k<\infty$: The dissipation range is characterized by a sharp drop in turbulent activity, due to the destructive action of viscosity and diffusion. In the troposphere, $l_s=k_s^{-1}$ is of the order of millimeters. We do not consider the dissipation range, since aperture and recorder smoothing effectively eliminate the contributions of such tiny blobs in the troposphere.[3]

Three turbulence theories have been developed which predict $S(k)$ in the inertial range:

a. Pressure Fluctuations

Fluctuations of the velocity field induce pressure variations according to Bernoulli's law. If variation of pressure, density, and water-vapor content are related linearly, the spectrum of dielectric fluctuations is given by [7],

$$S(k)=\left(\frac{v_0}{c}\right)^4 \frac{k_0^{4/3}}{k^{13/3}},\ (k>k_0) \tag{2.4.3}$$

where v_0 is the speed of the largest blobs and c the local speed of sound. Most writers now agree that this mechanism is unimportant for the lower troposphere [8].

b. Obukhov's Mixing Theory

In this theory, an external source is imagined to feed fluctuations into the spectrum at the largest size (k_0). Subsequent redistribution down the inertial range is attributed to turbulent convection, for which dimensional arguments give [9, 10, 11]

$$S(k)=\langle\Delta\epsilon^2\rangle \frac{k_0^{2/3}}{k^{11/3}},\ (k>k_0). \tag{2.4.4}$$

If the external source is identified only with the mixing of a gradient by the largest blobs, the intensity is given by

$$\langle\Delta\epsilon^2\rangle=k_0^{-2}\left(\frac{d\epsilon_0}{dh}\right)^2; \tag{2.4.5}$$

which is supplied at the input wavenumber k_0. Adiabatic expansions of the mixing eddies modify the ambient gradient in the stratosphere [6].

c. Mixing-in-Gradient Theory

This approach considers an initial refractive-index gradient and how it is changed by turbulent convection at *all* wavenumbers. The rotating turbulent eddies transfer parcels of water vapor from low to high points on the profile and vice versa, and establish irregular contrasts thereby. The mechanism for turbulent fluctuations appears explicitly in this theory at all wavenumbers and no external source is required. The inertial range is again amenable to unambiguous dimensional arguments [8, 12].

$$S(k)=\left(\frac{d\epsilon_0}{dh}\right)^2 \frac{1}{k^5},\qquad (k>k_0) \tag{2.4.6}$$

and the result is independent of turbulent parameters, except for restrictions on its range of validity. Adiabatic expansions of the eddies are again important in the stratosphere [6]. In the range $k>k_0$, we see that this model agrees with the spectrum computed from the Bessel space-correlation model indicated in table 1.

Both of the mixing theories predict that the intensity of turbulent fluctuations should be proportional to the local gradient of the inhomogeneous mean profile. A good correlation between the strength of scatter signals and the monthly mean refractive gradient has been observed experimentally on many paths [13]. Since strong scattered signals are associated with large dielectric fluctuations, this must be considered an important success for the gradient-dependent mixing mechanisms. On the other hand, observed absences of dielectric irregularities at inversion layers and other stable meteorological interfaces suggests that the mixing-in-gradient models may not be applicable to very sharp gradients.

[3] The dissipation is apparently quite important for ionospheric-scatter propagation at vhf.

2.5. Space-Time Correlations of Irregularities

The averaged product of two dielectric fluctuations taken at different positions and times is of considerable importance in studying the time variability of signal amplitude and phase for both line-of-sight and beyond-the-horizon propagation.

$$\langle \Delta\epsilon(\vec{r},t)\Delta\epsilon(\overrightarrow{r+R},t+\tau)\rangle = \langle \Delta\epsilon^2\rangle C(\vec{R},\tau) \quad (2.5.1)$$

Temporal variations of $\Delta\epsilon$ are associated with motion of the carrier medium, and two distinct types of atmospheric motion are important:

a. Drifting Convection

A prevailing wind with constant velocity $\vec{U}$ would bear a frozen turbulent structure along intact and would translate the entire structure a distance $\vec{U}\tau$ in a time τ. Since the atmosphere is assumed to be homogeneous, this would be just the same as measuring the space correlation between irregularities separated by a distance of $\vec{U}\tau$ at the same time.

b. Random Motion

The random motion associated with the continuous eddy-breakup process rearranges the turbulent irregularities in time. This self-motion effect is probably isotropic for small (k) scale irregularities. The random motion would be apparent even if there were no drift velocity, or if one rode along on the prevailing winds.

The spectrum method is particularly well-suited to describing the combined effect of these two motions, in that their effects can be explicitly separated in the integrand of the wavenumber integral representation of the space-time correlation [4, 14].

$$\langle \Delta\epsilon(\vec{r},t)\Delta\epsilon(\overrightarrow{r+R},t+\tau)\rangle = \frac{1}{8\pi^3}\int d^3k S(\vec{k}) e^{i\vec{k}\cdot(\overrightarrow{R+U\tau})}\eta(\vec{k},\tau) \quad (2.5.2)$$

For an isotropic spectrum of irregularities, this becomes:

$$\langle \Delta\epsilon(\vec{r},t)\Delta\epsilon(\overrightarrow{r+R},t+\tau)\rangle = \frac{1}{2\pi^2}\int_0^\infty dk k S(k)\frac{\sin(k|\overrightarrow{R+U\tau}|)}{|\overrightarrow{R+U\tau}|}\eta(k,\tau). \quad (2.5.3)$$

The function $\eta(k,\tau)$ describes the time autocorrelation of fluctuations contained in a fixed wavenumber interval k, and is unity for zero time displacement τ. Dimensional turbulence arguments indicate that $\eta(k,\tau)$ should depend only on the product[4] $V_0 k_0^{1/3} k^{2/3}\tau$ in the inertial subrange [15, 16], although no functional dependence has yet been established. However, detailed studies of the turbulent velocity field-time correlations over the entire spectrum indicate that the coefficient of τ varies from $k^{2/3}$ to k [17].

Convective motion is probably the dominant effect for line-of-sight propagation, since eddy speeds V_0 are usually much less than the prevailing wind speed U.[5] However, scatter propagation geometry discriminates against a horizontal drift velocity at the path midpoint and Doppler shifts introduced by the random motion can dominate.

3. Line-of-Sight Propagation

3.1. Introduction

Phase and amplitude instability on line-of-sight propagation paths is important in practice because of the limitations it imposes on interferometric tracking schemes and frequency synchronization operations. Before launching into detailed discussions of problems, it is well to examine the two principal geometrical situations shown in figure 3. These problems are idealized by omission of ground reflections, which are ignored since they do not contribute scintillations unless moving foliage or water surfaces are a major factor.

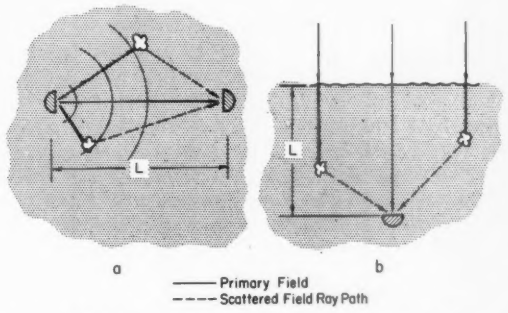


FIGURE 3.—*Typical ray paths for idealized line-of-sight problems: (a) Relay link immersed in an infinite turbulent medium, and (b) radio star viewed vertically through a turbulent mantel.*

a. Relay Link Problem

Receiver and transmitter are both immersed in an infinite turbulent region a distance L apart. The transmitter emits spherical waves and the receiver is also assumed to be omnidirectional.

b. Radio Star Problem

Plane waves are assumed to fall on a semi-infinite turbulent medium with the receiver located a distance L below the upper boundary. The turbulence may be graduated toward its upper boundary by

[4] V_0 and k_0 are the speed and wavenumber at the input stage of the turbulent velocity spectrum.

[5] Since the turbulent velocity field itself is a (small) partition of the kinetic energy in the laminar (drift) field.

using an inhomogeneous turbulence intensity $\langle \Delta\epsilon^2(r) \rangle$.

It is significant that the receiver always stands in the (tropospheric) turbulent medium for line-of-sight propagation, and blobs which are just on top of the receiver play an important role.

3.2. Signal Statistics and Vector Voltage Diagrams

The typical ray-path trajectories shown in figure 3 indicate that the field which falls on the receiver is composed of two terms: (a) The unperturbed primary field E_0 propagated along the line-of-sight, and (b) the incoherent superposition of waves scattered to the receiver by all volume elements in the scattering region V. The primary field induces a constant voltage A_0 in the receiver, which serves as a convenient phase reference. Because of the randomly varying multipath combinations, the total scattered wave induces statistical signal components $x(t)$ and $y(t)$ which are in and out-of-phase with respect to the constant vector A_0, as shown in the vector voltage diagrams of figure 4. Experiments measure the composite amplitude A and the relative phase excursions α about the phase reference A_0. A typical phase record is shown in figure 4, together with the vector voltage diagrams appropriate to two displaced times t_1 and t_2.

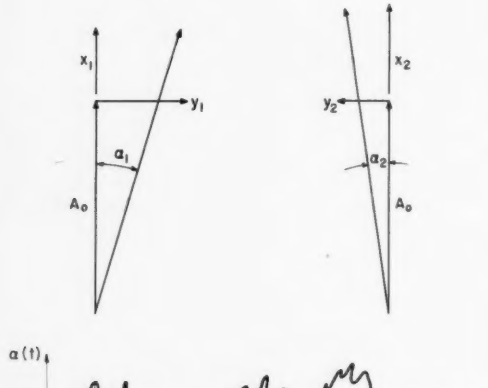


FIGURE 4.—*Typical phase record of randomly-varying line-of-sight signal showing the corresponding time-displaced vector voltage diagrams for a constant signal plus noise.*

The basic statistical problem is to estimate the probability distributions for amplitude A and phase α, and from these the relevant statistical estimates of the randomly varying quantities. Since the orthogonal signal components $x(t)$ and $y(t)$ are induced by the scattering contributions from many independent blobs, they may be characterized by a Gaussian distribution according to the Central Limit Theorem.

$$P[x,y]\,dx\,dy = \frac{dx\,dy}{2\pi\sigma_x\sigma_y}\frac{1}{\sqrt{1-\rho^2}}\exp-\left[\frac{\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-\frac{2\rho xy}{\sigma_x\sigma_y}}{2(1-\rho^2)}\right]. \tag{3.1.1}$$

The moment parameters,

$$\langle x^2\rangle=\sigma_x^2,\ \langle y^2\rangle=\sigma_y^2,\ \langle xy\rangle=\sigma_x\sigma_y\rho$$

should be determined from explicit propagation calculations. Previous statistical discussions have *assumed*,

$$\sigma_x^2=\sigma_y^2=\sigma^2, \text{ and } \rho=0, \tag{3.1.2}$$

in order to simplify the analysis. We should like to point out, however, that such assumptions are not in agreement with available propagation calculations, which suggest rather that

$$\sigma_x^2+\sigma_y^2=\text{constant},\ \rho \text{ finite}. \tag{3.1.3}$$

A new effort on the statistical problem is required for propagation work.

For present purposes, we simply review available results derived on the assumptions (3.1.2). If the (A_0) displaced cylindrical coordinates of figure 4 are inserted into eq (3.1.1), the joint amplitude-phase distribution becomes:

$$P(A,\alpha)dAd\alpha=\frac{AdAd\alpha}{2\pi\sigma^2}\exp -\left\{\frac{A^2+A_0^2-2AA_0\cos\alpha}{2\sigma^2}\right\} \tag{3.1.4}$$

The distribution of total voltage amplitudes A was established by Rice [18] from (3.1.4) by integrating the phase α from $-\pi$ to π.

$$P(A)dA=\frac{AdA}{\sigma_2}I_0\left(\frac{AA_0}{\sigma_2}\right)\exp-\left[\frac{A^2+A_0^2}{2\sigma^2}\right]. \tag{3.1.5}$$

From this expression one can calculate the amplitude moments.

$$\langle A\rangle=\sqrt{\frac{\pi}{2}}\,\sigma e^{-A_0^2/4\sigma^2}\left[\left(1+\frac{A_0^2}{2\sigma^2}\right)I_0\left(\frac{A_0^2}{4\sigma^2}\right)+\left(\frac{A_0^2}{2\sigma^2}\right)I_1\left(\frac{A_0^2}{4\sigma^2}\right)\right] \tag{3.1.6}$$

and

$$\langle A^2\rangle=A_0^2+2\sigma^2 \tag{3.1.7}$$

The probability that the amplitude exceeds a prescribed level a is obtained by integrating (3.1.5) from a to ∞. Numerical tables of this probability have been prepared [19] and are reproduced on Rayleigh graph paper in figure 5. The addition of a constant vector evidently increases the average signal

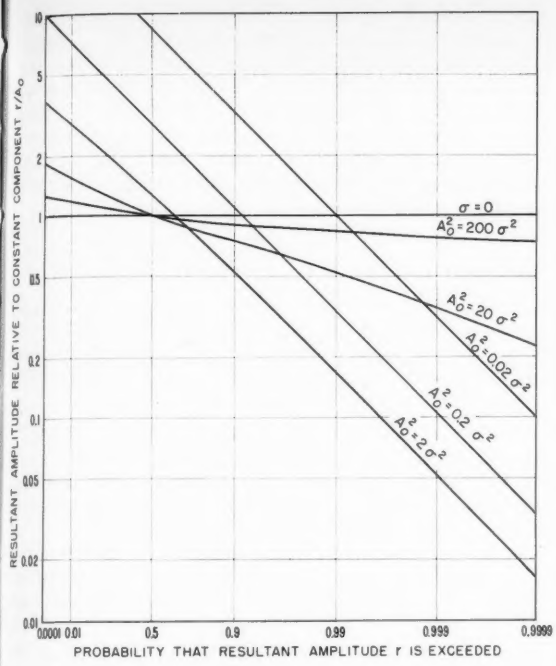

RESULTANT AMPLITUDE RELATIVE TO CONSTANT COMPONENT r/A0
PROBABILITY THAT RESULTANT AMPLITUDE r IS EXCEEDED
σ = 0
A0² = 200 σ²
A0² = 20 σ²
A0² = 0.02 σ²
A0² = 0.2 σ²
A0² = 2 σ²

FIGURE 5.—*Probability that the total amplitude exceeds a given level, measured relative to the amplitude of the constant vector A_0, from [19].*

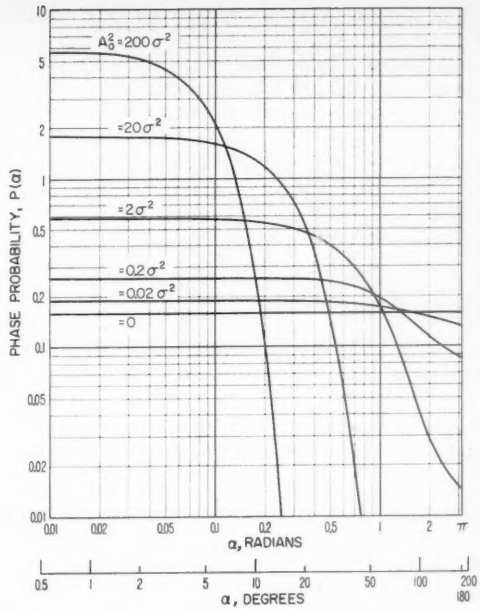


FIGURE 6.—*Distribution of phase variations about the reference voltage A_0 for line-of-sight propagation for various ratios of the steady signal to noise power.*

level distribution, but reduces the range of amplitude fading relative to A_0.

The distribution of phase is derived from eq (3.1.4) by integrating over all values of the amplitude A [20].

$$P(\alpha)d\alpha = \frac{d\alpha}{2\pi} e^{-A_0^2/2\sigma^2} \left\{ 1 + \sqrt{\frac{\pi}{2}} \frac{A_0}{\sigma} \cos\alpha \, e^{A_0^2 \cos^2\alpha/2\sigma^2} \times \left[1 + \operatorname{erf}\left(\frac{A_0 \cos\alpha}{\sqrt{2}\sigma} \right) \right] \right\} \quad (3.1.8)$$

This distribution is plotted in figure 6 and indicates that the phase is evenly distributed when the constant vector A_0 is very small compared with the rms multipath-scatter vector amplitude $\sqrt{2}\sigma$. When the constant vector is relatively large, Bremmer [20] proves that the phase is distributed about zero in a Gaussian manner,

$$\lim_{\sigma << A_0} P(\alpha) \to \frac{A_0}{\sigma\sqrt{2\pi}} \exp -\left(\frac{A_0^2}{2\sigma^2} \right) \alpha^2, \quad (3.1.9)$$

Norton et al. [21] have also calculated the probability that the phase exceeds a given angle for various values of $A_0/\sqrt{2}\sigma$. The first and second moments of phase deviations have been evaluated numerically [22] using the result (3.1.8), and are reproduced in figure 7 as functions of $2\sigma^2/A_0^2$.

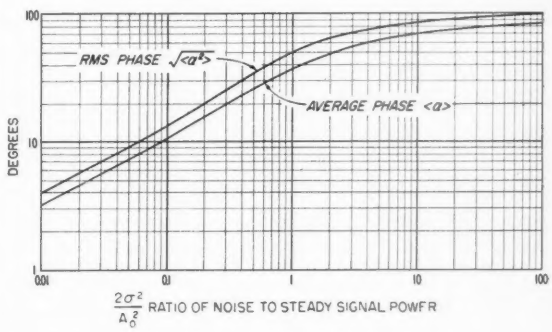


FIGURE 7.—*Average phase variation and root mean square phase variations about constant references voltage A_0 as functions of the amplitude ratio of signal-to-noise, from [22].*

Bremmer [20] has considered the probability distribution for time-displaced amplitudes (A_1, A_2) and phases (α_1, α_2) corresponding to figure 4. His development is based on special assumptions (see sec. 4.1.) for the time correlation between $x(t)$ and $y(t)$, but is quite interesting in that he establishes phase and phase-rate distributions for the same instant by manipulation of the probability density for time-displaced phases.

3.3. Electromagnetic Scattering by Irregularities

The stochastic dielectric constant (2.1.1) enters electromagnetic theory through the relation between the electric vector and displacement vector.

$$\vec{D}=\left[\epsilon_0+\Delta\epsilon(\vec{r},t)\right]\vec{E}. \tag{3.2.1}$$

When the magnetic field is eliminated from Maxwell's equations, the electric vector is found to satisfy [23],

$$\nabla^2\vec{E}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left[\epsilon\vec{E}\right]=-\vec{\nabla}\left[\frac{1}{\epsilon}\vec{E}\cdot\vec{\nabla}\epsilon\right]. \tag{3.2.2}$$

Time derivatives here operate principally on $\vec{E}$, since the field oscillates much more rapidly than the medium. The vector components of $\vec{E}$ are mixed in the right-hand term and give rise to depolarization effects. The magnitude of this term is small for line-of-sight microwave propagation and is usually omitted. (See, however, references [24] and [25].) With a harmonic time dependence ($k_f=\omega/c=2\pi/\lambda$), one finds

$$[\nabla^2+k_f^2(\epsilon_0+\Delta\epsilon)]\vec{E}=0, \tag{3.2.3}$$

which is the basic propagation equation.

It is not possible to solve (3.2.3) exactly, since $\Delta\epsilon(\vec{r},t)$ is an unknown (stochastic) function of position. The Born approximation is widely used to describe single scattering by such fluctuations. This is essentially a three-dimensional iterative solution of the wave equation, which gives the total field in terms of the unperturbed wave $\vec{E}_0(\vec{r})$ as

$$\vec{E}(\vec{R})=\vec{E}_0(\vec{R})-k_f^2\int_V d^3r G(\vec{R},\vec{r})\Delta\epsilon(\vec{r},t)\vec{E}_0(\vec{r})\sin\chi \tag{3.2.4}$$

where V is the volume of irregularities illuminated by both the transmitter and receiver, χ is the angle between $\vec{E}_0$ and the Poynting vector of the reradiated wave (see fig. 21) and is nearly 90 degrees for line-of-sight propagation. $G(\vec{R},\vec{r})$ is the free-space spherical Green's function connecting the scattering point $\vec{r}$ and receiver $\vec{R}$.

$$G(\vec{R},\vec{r})=\frac{e^{ik_f|\vec{R}-\vec{r}|}}{4\pi|\vec{R}-\vec{r}|} \tag{3.2.5}$$

The first term in (3.2.4) represents the transmitted wave and the integral expression is the scattered wave. The total field induces voltage components in the receiver, as shown in figure 4. The primary field E_0 sets the phase reference A_0. The integral term gives rise to the in- and out-of-phase random signal components $x(t)$ and $y(t)$, corresponding to the real and imaginary parts of the integral term in (3.2.4) respectively. The receiver measures the total voltage A and its phase α. It is inherent in the Born approximation that E_s must be small compared with the mean signal E_0, so that instantaneous fluctuations of the total amplitude and phase are (consistently) approximated by:

$$\delta A(t)\simeq x(t)=k_f^2\int_V d^3r\Delta\epsilon(\vec{r},t)\ \mathrm{Re}\ \{G(\vec{R},\vec{r})E_0(\vec{r})\}, \tag{3.2.6}$$

$$\alpha(t)\simeq\frac{y(t)}{A_0}=\frac{k_f^2}{E_0(\vec{R})}\int_V d^3r\Delta\epsilon(\vec{r},t)\ \mathrm{Im}\ \{G(\vec{R},\vec{r})E_0(\vec{r})\}. \tag{3.2.7}$$

The variances of phase and amplitude fluctuations thus depend upon six-fold integrals of the space correlation function (2.1.3)

It is common practice to use the geometrical optics approximation to describe line-of-sight propagation. This is equivalent to a one-dimensional (WKB) solution of the wave eq (3.2.3), which recognizes only the random speeding-up and slowing-down of the wave as it propagates along the ray path (s). The phase of the vector voltage diagram in figure 4 is given directly by this method,

$$\alpha=\frac{\pi}{\lambda}\int_0^L ds\Delta\epsilon(s,t), \tag{3.2.8}$$

and is to be computed along the nominal (rectilinear) ray path s. This expression is always a good approximation to the volume integral (3.2.7), and it is generally used to study phase variations because of its simplicity. Geometrical optics also predicts amplitude variations of the received signal. These are due physically to bunching and diverging of the energy-bearing rays by random refractive bending;

$$\frac{\delta A}{A_0}=\frac{\lambda}{2\pi}\int_0^L ds\left[\frac{\partial^2\alpha}{\partial x^2}+\frac{\partial^2\alpha}{\partial y^2}\right] \tag{3.2.9}$$

where α is given by expression (3.2.8). This expression is usually not a good approximation to the more precise wave theory result (3.2.6).

The geometrical optics approximation is a special case of the single scattering wave theory of eq (3.2.4) [4, 5, 26]. To understand this correspondence, note that the scale length of the irregularities l_0 is ordinarily much larger than the wavelength of the radiation. The blobs act like very shallow lenses and project almost all of the scattered radiation in a cone of angular opening λ/l_0 about the line-of-sight. The effective scattering volume over a path length L, therefore, has a maximum width $L\lambda/l_0$ as shown in figure 8. If this width is much less than the average blob size, the scattering volume consists of cylindrical plugs "cored" from successive blobs as in figure 8a. The propagation is then substantially one dimensional

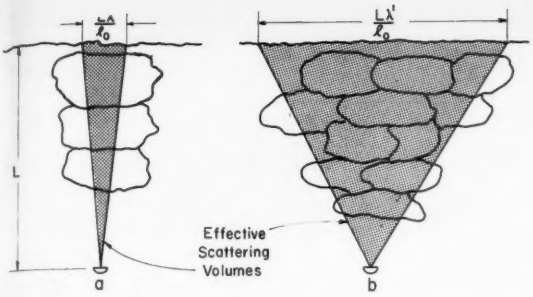


FIGURE 8. *Pictorial representation of the effective scatter volume for line-of-sight radio star problem: (a) Fresnel scattering, and (b) Fraunhofer scattering.*

and random variation of the phase speed of the wave is the measured effect. This corresponds to Fresnel scattering and is correctly treated by geometrical or ray optics. The opposite extreme occurs when $L\lambda/\pi l_0 > > l_0$; the receiver sees many whole blobs as in figure 8b. The problem is then necessarily three dimensional and requires the wave solution (3.2.4), corresponding to Fraunhofer scattering. The scattering parameter

$$\gamma = \frac{L\lambda}{l_0^2} \tag{3.2.10}$$

is evidently the important variable in these discriminations. One finds that the wave theory expressions (3.2.6) and (3.2.7) reduce exactly to the geometrical optical results (3.2.9) and (3.2.8) if γ is less than unity. If γ is greater than (or near) unity, however, one must use the full-wave theory to compute amplitude scintillations.

The single-scattering approximation (3.2.4) breaks down at large distances L and/or high frequencies, since the convergence of the Born series is related to the smallness of the mean square phase expression [5],

$$<\alpha^2> \cong \pi^2 <\Delta\epsilon^2> \frac{Ll_0}{\lambda^2}. \tag{3.2.11}$$

As the wave progresses deeper into the turbulence, scattering out of the main beam gradually dissipates the primary signal A_0 in figure 4 and one is left only with the correspondingly increased random field E_s. The method of Rytov [27] has been used to describe this "phase creep" zone, where the amplitude and phase fluctuations are not small.

$$\vec{E}(\vec{R}) = \vec{E_0}(\vec{R}) \exp - \left[\frac{k_1^2 \rightarrow}{\vec{E_0}(\vec{R})} \cdot \int d^3 r \vec{G}(\vec{R},\vec{r}) \Delta\epsilon(\vec{r},t) \vec{E_0}(\vec{r}) \right] \tag{3.2.12}$$

This expression reduces to (3.2.4) in the limit of small signal variations, but it is not clear how this method accounts for multiple scatterings. It is significant that the integral term in the exponent of Rytov's expression is identical to the first Born term in eq (3.2.4). Difficult propagation calculations of the real and imaginary parts of this integral term computed for single scattering theory can thus be used in calculations based on (3.2.12), once the appropriate signal-statistical description is established.

Feinstein [28] used the Kirchhoff-Huygens approximation to study propagation of scalar waves through a succession of uncorrelated random slabs. His use of the second order phase approximation is equivalent to tracing the phase front along various ray paths, or neglecting diffraction over the slab thickness. The result can be expressed as a power series in $<\Delta\epsilon^2>$, the first term of which is the geometrical-optics approximation.

3.4. Propagation Calculations Based on Geometrical Optics

The foregoing discussion presented general expressions for the phase and amplitude scintillations of waves which are propagated through random media. To compare theory and experiment, it is necessary to combine these propagation expressions with the descriptions of turbulent media given in part II. The results depend both on the propagation geometry and the model chosen to represent the irregularities.

The calculation of phase variations is usually based on the geometrical optics approximation. The variance of phase fluctuations $<\alpha^2>$ about the constant reference vector A_0 can be computed by simply squaring and averaging the integral expression (3.2.8). Using any of the correlation models described in section 2.2 gives the following simple result,

$$<\alpha^2> = q\pi^2 <\Delta\epsilon^2> \frac{Ll_0}{\lambda^2}, \tag{3.3.1}$$

where L is the line-of-sight path length and λ the wavelength of the radiation employed. q is a constant of order unity which has been evaluated for each of the four correlation models. References to such detailed calculations are summarized in table 2, together with references to other phase and angle-of-arrival quantities. The reader should consult these references for specific results.

The wavelength dependence of expression (3. 3. 1) has been confirmed experimentally by simultaneous measurements [33, 34] of phase shifts on three frequencies, 100, 1,000, and 10,000 Mc, on line-of-sight paths to mountain tops. The path length (L) dependence is difficult to check experimentally, since paths of different lengths are apt to go through regions of the troposphere which do not have the same level of turbulent activity. Attempts to correlate phase variations with airborne and ground-based refractometer measurements of $<\Delta\epsilon^2>$ and l_o in the propagation paths have been moderately successful [3, 30].

Such results are conveniently summarized with the

TABLE 2. *References for line-of-sight calculations of phase and amplitude fluctuations using geometrical optics for four space correlation models of the irregularities*

Calculated quantity \ Space correlation model	Bessel $\rho K_1(\rho)$	Exponential $e^{-\rho}$	Gaussian $e^{-\rho^2}$	Cauchy $[1+\rho^2]^{-2}$
Mean square phase	30	29, 30	29, 30	29, 30
Finite path corrections to $<\alpha^2>$	30	30	30	30
Aperature smoothing corrections to $<\alpha^2>$	31	31	31	----------
Finite data sample corrections to $<\alpha^2>$	----------	----------	----------	----------
Parallel path space correlation	----------	29	29	29
Rotated path space correlation	30	29	29	29
Time correlation of phase	30	29, 30	29, 30	29, 30
Power spectrum of phase	30	29, 30	29, 30	29, 30
Phase fading rate	30	Diverges	32	32
Number of phase maxima	Diverges	Diverges	----------	----------
Mean square angle-of-arrival	----------	Diverges	29	29
Time correlation of angle-of-arrival	----------	29	29	29
Spectrum of angle-of-arrival	----------	29	29	29
Mean square amplitude variance	Diverges	Diverges	----------	----------

spectrum method, suggested in section 2.3, where a model choice is delayed until after the *propagation* calculations have been performed. In terms of the blob-size spectrum $S(k)$, the mean square phase fluctuation becomes [4]:

$$<\alpha^2>=\frac{\pi L}{2\lambda^2}\int_0^\infty dk k S(k), \qquad (3.3.2)$$

which exhibits directly the dependence of the measured ($<\alpha^2>$) on the controlled variables L and λ. The spectrum method is particularly advantageous when computing corrections to this expression due to aperature smoothing, finite data sample lengths (T) and finite propagation path lengths (L). These effects can be analyzed exactly within the framework of geometrical optics, and produce appropriate factors in the wavenumber integration which filter the spectrum $S(k)$ preferentially. Analyzing the individual effects separately and multiplying them serially one finds [4].

$$<\alpha^2>=\frac{\pi L}{2\lambda^2}\int_0^\infty dk k S(k)\left[\frac{2J_1(ka)}{ka}\right]^2 \times\left[\frac{2}{\pi}\left\{Si(kL)-\frac{1-\cos(kL)}{kL}\right\}\right] \times\left[1-\frac{2}{kUT}\int_0^{kUT}dz J_0(z)+2\frac{J_1(kUT)}{kUT}\right]. \qquad (3.3.3)$$

This composite spectrum filtering is illustrated in figure 9. The first factor accounts for the masking of small blob (large wavenumber) contributions to $<\alpha^2>$ due to smoothing by a receiving aerial of

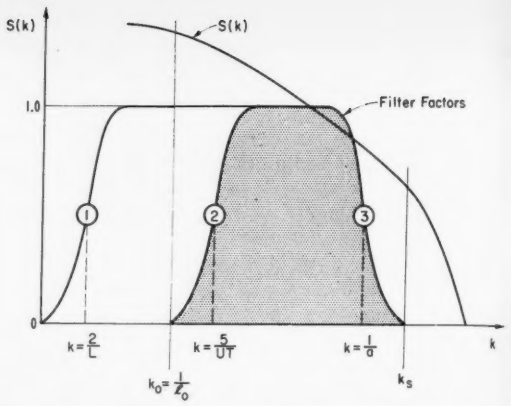


FIGURE 9. *Typical spectrum showing preferential filtering factors: (1) finite path length, (2) finite data sample length, and (3) aperture smoothing corrections.*

radius a. The second factor describes finite path length correlations, and indicates that blobs larger than the total path length L cannot contribute effectively to signal fluctuations. The last factor describes the effect of a finite data sample length T in suppressing very low-frequency scintillations, assuming that they are due only to a convective speed U. This predicts that the rms phase variation should increase with sample length T, which is observed experimentally [3]; although nonstationary trends could give the same result. References to explicit calculations of the individual effects for various spectral models (i.e., correlation functions) are given in table 2.

The space correlation between phase variations measured on adjacent propagation paths has been calculated for a variety of turbulence models (see table 2) for both the rotated line-of-sight and parallel path configurations indicated in figure 10.

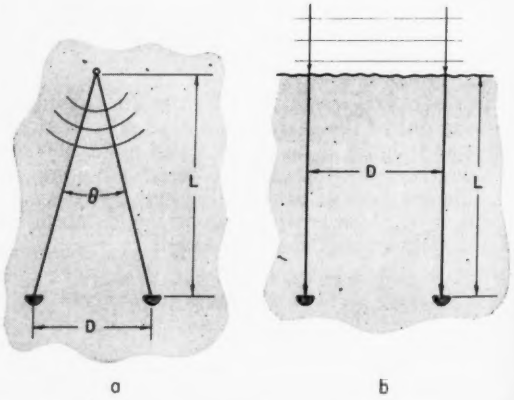


FIGURE 10. *Geometry for space correlation of phase experiments: (a) Rotated paths with common transmitter and (b) parallel rays from source of infinity.*

The spectrum method summarizes these results in succinct form [4].
Parallel rays:

$$\langle\alpha(x)\alpha(x+D)\rangle=\frac{\pi L}{2\lambda^2}\int_0^\infty dkkS(k)J_0(kD), \quad (3.3.4)$$

Rotated rays:

$$\langle\alpha(\phi)\alpha(\phi+\theta)\rangle=\frac{\pi L}{2\lambda^2}\int_0^\infty dkkS(k)\,\frac{1}{kL\theta}\int_0^{kL\theta}duJ_0(u). \quad (3.3.5)$$

These filter factors essentially eliminate the contribution of small blobs (large wavenumbers), which are not large enough to affect both paths simultaneously.

It is found that the Bessel model gives good agreement with spaced receiver experiments [33,30] for small baseline distances. At very large separations, however, the experimental correlation falls more slowly than the theory predicts. This is tentatively ascribed to lack of spatial homogeneity and/or temporal stationarity. Since the correlation at great distances is due to the joint influence of very large blobs on the two propagation paths, the phase correlation should depend on the large R behavior of $C(R)$, which is not tied to physical theories, as noted in section 2.4.

The random variation of single path phase records with time is due to the action of drifting convection and turbulent self motion. The space-time correlation of dielectric fluctuations (2.5.3) permits one to calculate the time correlation of phase records quite generally for an arbitrary spectrum. If the convective wind blows normal to the propagation path [4],

$$\langle\alpha(t)\alpha(t+\tau)\rangle = \frac{\pi L}{2\lambda^2}\int_0^\infty dkkS(k)J_0[kU\tau]\eta[k,\tau]. \quad (3.3.6)$$

The effect of self-motion is usually discarded (i.e., $\eta=1$), since the turbulent speed (V_0) is much less than the drift speed U.

The experimental data on phase scintillations is usually expressed in terms of frequency spectra of the phase records. For drift motion alone [5],

$$W(f)=\frac{\pi L}{2\lambda^2}\int_{\frac{2\pi f}{U}}^\infty \frac{dkkS(k)}{[k^2U^2-(2\pi f)^2]^{1/2}}. \quad (3.3.7)$$

The mixing-in-gradient model (2.4.6) predicts that the frequency spectrum $W(f)$ should vary as f^{-4}, while the Obukhov model (2.4.4) predicts an $f^{-2.7}$ variation. The most recent data [3] indicates a frequency variation of $f^{-2.8}$ from 1 cycle per hour to 10 cps for very low-level paths. This data argues strongly for the Obukhov Mixing Theory, unless self-motion effects play a dominant role.

The average number of mean (zero) crossings and maxima can be expressed in terms of ratios of moments of the frequency spectrum $W(f)$, which in turn can be related to moments of the turbulent spectrum $S(k)$[4]. Filter factors at the high-wavenumber end, such as that provided by aperture smoothing, are especially important if such results are to be compared with experimental data.

Tracking devices which measure bearing angles ordinarily measure the normal to the instantaneous phase front by interferometric means. Angle-of-arrival errors are thus related to the space coherence (correlation) of phase scintillations along this phase front. The mean square angle of arrival is given by [4]

$$\langle\delta\theta^2\rangle=\frac{L}{16\pi}\int_0^\infty dkk^3S(k), \quad (3.3.8)$$

and has been evaluated explicitly in the references of table 2. The autocorrelation of $\delta\theta(t)$ and the corresponding frequency spectrum have also been discussed in general [4] and specific terms [29].

Amplitude fluctuations have not received as much attention as phase (see however, ref [36] and [37]) since $\delta A/A_0$ must be small compared with $(\alpha)_{\text{rms}}$ for the geometrical optics expression (3.2.9) to be valid. In terms of the spectrum, the percentile amplitude fluctuation is given by

$$\left\langle\left|\frac{\delta A}{A_0}\right|^2\right\rangle\simeq L^3\int_0^\infty dkk^5S(k), \quad (3.3.9)$$

which depends critically on the behavior of the correlation model near the origin (i. e., $k=\infty$). Only the Gaussian and Cauchy correlation models predict finite results, viz,

$$\left\langle\left|\frac{\delta A}{A_0}\right|^2\right\rangle=g\langle\Delta\epsilon^2\rangle\,\frac{L^3}{l_0^3}, \quad (3.3.10)$$

where g is a constant of order unity. This distance variation and absence of a frequency dependence does not agree with most experiments [33, 38], so one must seek a more satisfactory solution in the wave theories.

3.5. Propagation Calculations Based on Wave Theory

When the scattering process is necessarily three-dimensional in the sense of figure 8, one must use the Born expressions (3.2.6) and (3.2.7) for line-of-sight amplitude and phase. The variance of phase and amplitude depend upon six-fold integrals of the space correlation function (2.1.3). Such calculations have been performed primarily with the Gaussian function, which has important analytical-simplification properties, although even then one must exploit approximations.

The spectrum method has simplified the evaluation of these wave theory expressions enormously. Using the method described previously [4], one can establish the following general expressions for the radio-star problem:

$$\left\langle \alpha^2 \right\rangle = \frac{\pi L}{2\lambda^2}\int_0^\infty dk k S(k)\left[1-\Phi\left(\frac{k^2 L\lambda}{2\pi}\right)\right],$$

$$\left\langle \left|\frac{\delta A}{A_0}\right|^2 \right\rangle = \frac{\pi L}{2\lambda^2}\int_0^\infty dk k S(k)\Phi\left(\frac{k^2 L\lambda}{2\pi}\right),$$

$$\left\langle \left|\alpha \frac{\delta A}{A_0}\right| \right\rangle = \frac{\pi L}{2\lambda^2}\int_0^\infty dk k S(k)\Psi\left(\frac{k^2 L\lambda}{2\pi}\right), \quad (3.4.1)$$

where the fundamental scattering filter functions are defined by

$$\Phi(x) = \frac{x^2}{2(1+x^2)} + 0\left(\frac{1}{kL}\right), \quad (3.4.2)$$

$$\Psi(x) = \frac{x}{2(1+x^2)} + 0\left(\frac{1}{kL}\right). \quad (3.4.3)$$

The corresponding filter factors in the spectral integrals of (3.4.1) are plotted together in figure 11.

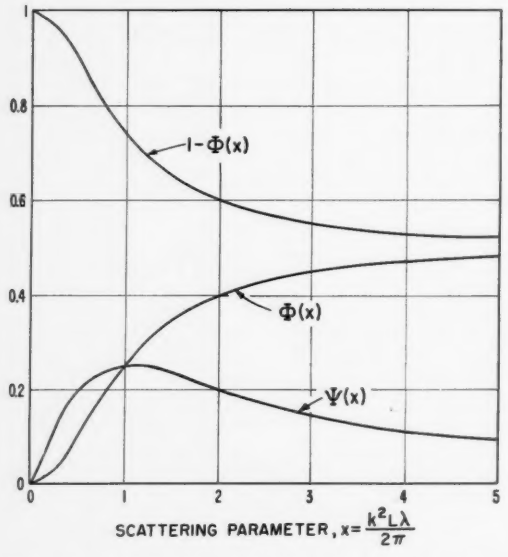


FIGURE 11. *Fundamental scattering filter factors for wave theory treatments of line-of-sight propagation.*

These functions depend upon the scattering parameter $\gamma = k^2 L\lambda$ for each blob-size group in the distribution described by $S(k)$. Those blobs in $S(k)$ which scatter in the Fraunhofer zone ($\gamma \gg 1$) produce equal phase and amplitude variations, each one-half as great as the ray theory result (3.3.2). The Fresnel regime ($\gamma \ll 1$) gives the geometrical optics phase result (3.3.2) directly. Using the small argument behavior $\Phi(x) \simeq \frac{x^2}{2}$, the geometrical optics result for amplitude fluctuations (3.3.9) is regained.

Explicit expressions for the phase and amplitude correlations can be established by introducing a particular spectrum model from table 1 into (3.4.1) and performing the single remaining wavenumber integration. Barrows [35] has evaluated these integrals for the Exponential, Gaussian, and Bessel models, and finds

$$\langle \alpha^2 \rangle = \frac{2\pi^2 \langle \Delta\epsilon^2 \rangle L l_0}{\lambda^2} F\left(\frac{L\lambda}{2\pi l_0^2}\right)$$

$$\left\langle \left|\frac{\delta A}{A_0}\right|^2 \right\rangle = \frac{2\pi^2 \langle \Delta\epsilon^2 \rangle L l_0}{\lambda^2}\left[1 - F\left(\frac{L\lambda}{2\pi l_0^2}\right)\right]$$

$$\left\langle \alpha \frac{\delta A}{A_0} \right\rangle = \frac{2\pi^2 \langle \Delta\epsilon^2 \rangle L l_0}{\lambda^2} G\left(\frac{L\lambda}{2\pi l_0^2}\right) \quad (3.4.4)$$

where the scattering parameter-dependent functions have approximately the same form as the filtering functions $\Phi(x)$ and $\Phi(x)$. The results obtained in this way are substantially the same as those obtained by Fannin [42], Mintzer [6], Obukhov [39], and Chernov [40] by the more laborious direct calculation of six-fold volume integrals. Since their results differ in details of propagation geometry, space correlation models, and scattering theory, we have constructed table 3 to summarize the essential assumptions and results of each paper.

The derivation of expressions (3.4.1) given in reference [4] does not lend itself to calculating space and time correlation effects. A new method was devised which combines the integral expressions (3.2.6) and (3.2.7) with the original representation

TABLE 3. *Explicit calculations of phase and amplitude fluctuation quantities using three dimensional wave theory*

Author	Reference	Scattering theory	Propagation model	Space correlation model	Phase variation	Amplitude variance	Phase-amplitude correlation	Space correlation	Time correlation	Others
Obukhov	39	Rytov	Radio star	Gauss	✓	✓				
Chernov	40	do	do	do	✓	✓	✓	✓		
Tatarski	41		do	Exp		✓				
Mintzer	6	Born	Relay link	Gauss		✓			✓	✓
Fannin	42	do	do	do	✓	✓		✓	✓	
Ellison	26	do	Radio star	do		✓		✓		
Wheelon	23	do	do	do	✓	✓		✓		
Barrows	35	do	do	Bessel Gauss	✓	✓	✓			
Wheelon	5	do	do	General	✓	✓				

(2.3.1) of the space correlation function, rather than the specialized form of eq (2.3.2). For example,

$$\langle\alpha^2\rangle = k_f^4 \int_V d^3r \,\mathrm{Im}[G(\vec{R},\vec{r})\,E_0(\vec{r})] \int_V d^3r' \,\mathrm{Im}[G(\vec{R},\vec{r}')\,E_0(\vec{r}')]$$

$$\times \frac{1}{8\pi^3} \int d^3k S(k) \exp i\vec{k}\cdot(\vec{r}-\vec{r}')$$

$$= \frac{k_f^4}{8\pi^3} \int d^3k S(k) \left| \int_V d^3r e^{i\vec{k}\cdot\vec{r}} \,\mathrm{Im}\,[G(\vec{R},\vec{r})\,E_0(\vec{r})] \right|^2 \tag{3.4.5}$$

Barrows [35] has found that the volume integral can be evaluated rigorously for either the radio star or relay link problem. In principle one can include the influence of ground reflections indicated in figure 12, by simply redefining the Green's function $G(\vec{R},\vec{r})$ and incident field $E_0(\vec{r})$ in eq (3.4.5) to satisfy the boundary conditions at the earth's surface. This can be accomplished analytically for a perfectly conducting earth by simply adding image sources and image scattering blobs to represent the ground-bounce paths.

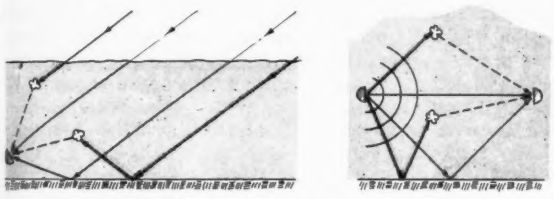


FIGURE 12. *Typical ray paths for line-of-sight propagation including ground reflected effects: (a) Relay link & (b) radio star problems.*

The new representation (3.4.5) is well adapted to describing time autocorrelations of the phase. Using the space-time correlation spectral representation (2.5.2), we find

$$\langle\alpha(t)\alpha(t+\tau)\rangle = \frac{k_f^4}{8\pi^3} \int d^3k S(k)\eta(k,\tau) e^{-i\vec{k}\cdot\vec{U}\tau}$$

$$\times \left| \int_V d^3r e^{i\vec{k}\cdot\vec{r}} \,\mathrm{Im}\,[G(\vec{R},\vec{r})\,E_0(\vec{r})] \right|^2 \tag{3.4.6}$$

which uses precisely the same geometrical propagation integrals required for the phase variance calculation in (3.4.5). Space correlations follow directly if one replaces the drift displacement $\vec{U}\tau$ with the receiver separation $\vec{D}$, providing the appropriate displaced coordinate system indicated by figure 13 is employed. If the wavenumber integration is performed in polar coordinates, it is found that the geometrical integrals do not depend on the (k) azimuth angle, and therefore,

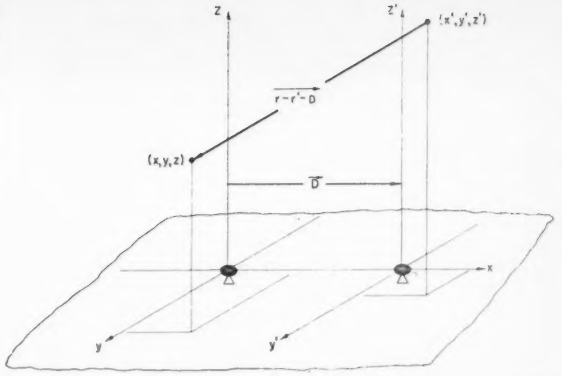


FIGURE 13. *Coordinates for calculation of phase correlation between adjacent receivers.*

$$\langle\alpha(x)\alpha(x+D)\rangle = \frac{k_f^4}{16\pi^2} \int_0^\infty dk k^2 S(k) \int_0^\pi d\theta \sin\theta J_0(kD\sin\theta)$$

$$\times \left| \int_V d^3r e^{i\vec{k}\cdot\vec{r}} \,\mathrm{Im}\,[G(\vec{R},\vec{r})\,E_0(\vec{r})] \right|^2 . \tag{3.4.7}$$

The effect of aperture smoothing with a circular parabolic dish of radius a can be evaluated by the same integration technique used with the less accurate geometrical optics solution [4, 31].

$$\langle\alpha_T^2\rangle = \left\langle \left| \frac{1}{A} \iint_A d\sigma\alpha(\sigma) \right|^2 \right\rangle$$

$$= \frac{k_f^4}{16\pi^2} \int_0^\infty dk k^2 S(k) \int_0^\pi d\theta \sin\theta \left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta} \right]^2$$

$$\times \left| \int_V d^3r e^{i\vec{k}\cdot\vec{r}} \,\mathrm{Im}\,[G(\vec{R},\vec{r})\,E_0(\vec{r})] \right|^2 \tag{3.4.8}$$

Even the difficult question of polarization dependence can be discussed conveniently in this framework, if a tensor Green's function is used in the Born approximation [24, 25]. In this way one can build up theoretical estimates of considerable interest to measurement programs. The method keeps the geometrical and turbulence problems separated until the last possible step, so as to allow each to develop as fully as possible.

3.6. Physical Interpretation of Scattering

It is valuable to pause briefly to re-examine the mathematical expressions (3.2.4) introduced by the Born approximation. Figure 14a indicates a typical singly-scattered ray trajectory for the radio-star problem. The scattered ray path concept may be identified with terms in the integrand of (3.2.4). The incident plane wave $\vec{E}_0(\vec{r})$ propagates rectilinearly to the scattering volume element d^3r. The dielectric fluctuations $\Delta\epsilon(\vec{r},t)$ scatters a small fraction

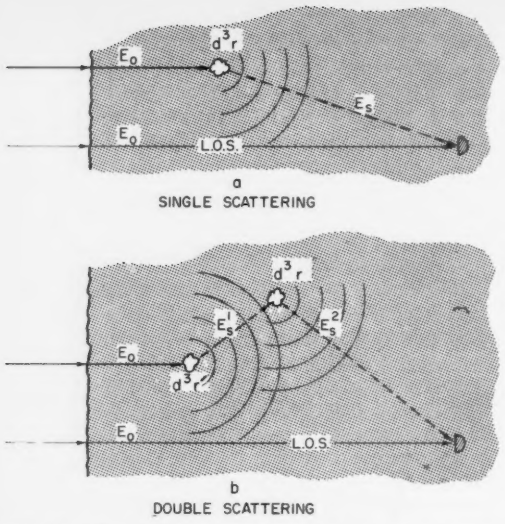


FIGURE 14. *Ray paths for single and double scattering of plane wave by a semi-infinite region.*

of the wave $\vec{E}_0(\vec{r})$ in all directions. (This reradiation is strongly concentrated in the forward direction, since the refractive variations act like very weak, overlapping lenses with radii of curvature l_0 much larger than the radiation wave length λ). A portion of the scattered energy is propogated along the line connecting the receiver $\vec{R}$ and scatterer $\vec{r}$. The propagation of this component is described by the Green's function (3.2.5), which characterizes the spherical expansion of the (unit) scattered wave. The total scattered field is simply the summation of all such multipath contributions, as given by the integration (3.2.4). These scattered waves are out of phase with respect to the primary signal $\vec{E}_0$ propagated along the line-of-sight, in virtue of their longer path lengths. Random fluctuations of the strength of these scattered components (i.e., $\Delta\epsilon$) give rise to the characteristic fading suggested by figure 4.

The typical ray path for double scattering processes indicated in figure 14b can be compared with the second Born approximation

$$E(\vec{R}) = E_0(\vec{R}) - k_f^2 \int_V d^3r G(\vec{R},\vec{r}) \Delta\epsilon(\vec{r},t) E_0(\vec{r}) + k_f^4 \int_V d^3r \int_V d^3r' G(\vec{R},\vec{r}) \Delta\epsilon(\vec{r},t) G(\vec{r},\vec{r}') \Delta\epsilon(\vec{r}',t) E_0(\vec{r}') \tag{3.5.1}$$

The double integral term has the following meaning: an incident wave $E_0(\vec{r}')$ falling on the first scattering element d^3r' is scattered by $\Delta\epsilon(\vec{r}',t)$, and reradiated to $\vec{r}$ by $G(\vec{r},\vec{r}')$. The second element $\Delta\epsilon(\vec{r},t)$ scatters the wave again and reradiates the result to the receiver via $G(\vec{R},\vec{r})$. The double integration simply sums the contributions over all possible pairs of scattering elements.

This tracing of rays between scattering events is similar to Feynman's diagramatic technique for describing the space-time history of particles in Quantum Electrodynamics [43]. In treating the dielectric fluctuations $\Delta\epsilon$ as a sequence of scattering events, however, one must give up the concept of a continuously refracting medium. Phase fluctuations are no longer regarded as the result of random changes in the velocity of propagation along the line-of-sight path. Rather, phase instability is looked upon as the random composition of many multipath components with the primary signal (see ref [44]).

Our description has been something less than accurate thus far in that expansions like (3.5.1) imply more power at the observation point $\vec{R}$ than would have appeared in the absence of refractive fluctuations $|\vec{E}_0|^2$. We have evidently overlooked scattering out of the primary beam on the line-of-sight ray path to the receiver. A scattering event robs the beam of some energy at each point along this path. Such losses are just compensated by energy which is projected back into the receiver by off-axis blobs. Inspection of figure 15 shows that for each point to which the line point scatters energy out of the beam, there is a corresponding point on the wave normal which projects an equal amount of energy back into the receiver at the same angle. If the receiver is isotropic, the two effects just cancel.

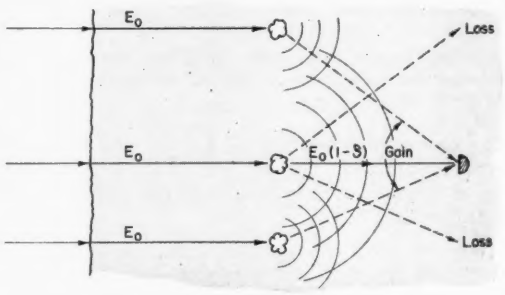


FIGURE 15. *Singly-scattered ray paths for radio star problem indicating power balance between scattering losses out of the beam and projection of energy into the receiver by off-axis blobs.*

Energy can re-enter the primary ray only by a double scattering process, which changes its direction back to the original line-of-sight. This means that the power balance must be maintained at each level of iteration, and is related to the unitary relationships which must hold between the various terms in the Born series. It is important to observe that such conservation cannot be demanded of the fluctuating field strength, but only of the averaged power.

4. Scatter Propagation

The transmission of telephonic and even television signals well beyond the optical horizon at UHF frequencies is such an important component of world-wide communications today that it requires no historical introduction. Suffice it to say that weak but dependable communication signals are received out to distances of 1,000 km if sufficient transmitter power and receiver sensitivity are employed. The signal envelope is observed to fade in an irregular manner. This suggested that the propagation might be due to radiowave scattering by the refractive irregularities (turbulence) which were known to exist in the troposphere. The mathematical theory of this scattering was developed by Pekeris [45] and Booker and Gordon [46], using the single scattering approximation.

The refractive scattering explanation of microwave propagation beyond the horizon is illustrated by figure 16. Turbulent blobs in the common volume of the receiver and transmitter are imagined to scatter minute fractions of the line-of-sight fields into the receiver. The case of omnidirectional aerials is shown in figure 16a, and indicates that the common volume is defined by the tangent planes to the transmitter and receiver. The incoherent addition of signal contributions throughout this volume gives rise to random orthogonal voltage components in the receiver, as suggested by the drawing. The narrow-beam case, shown in figure 16b, indicates that the corresponding scattering volume is controlled by the antenna patterns.

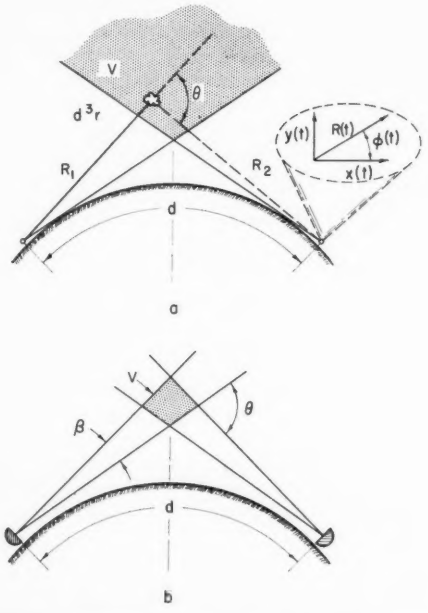


FIGURE 16. *Geometry for scatter propagation beyond the horizon: (a) Broad beam aerials, and (b) narrow beam aerials.*

A substantial amount of reliable experimental data was first published in 1955 [47]. The overall experimental situation at that time was ably summarized by Bullington's review paper [48]. Additional results have been published sporadically in succeeding years, and it is somewhat difficult to establish the precise state of affairs at this time. An up-to-date summary of theoretical research on the subject in the United States was presented to URSI in 1957 by Staras and the author [49]. The following section extends that report by describing a considerable amount of unpublished work. We also attempt to correlate the theoretical predictions with experimental results where possible.

4.1. Signal Statistics and Vector Voltage Diagram

The electromagnetic field received beyond the horizon via the scatter mode induces a complex voltage

$$x(t)+iy(t)=R(t)\ e^{i\phi(t)} \tag{4.1.1}$$

in the receiver, as shown in figure 16. The orthogonal signal components x and y result from the in and out-of-phase incoherent addition of waves scattered from many independent blobs in the common volume. Application of the Central Limit Theorem to this scattering process by many blobs ensures that $x(t)$ and $y(t)$ are distributed in a Gaussian manner. The components change randomly with time because of the changing phase relationships of the scattering contributions from randomly moving blobs. The processes $x(t)$ and $y(t)$ are probably stationary over brief observation periods, but do exhibit diurnal and seasonal variations which are quite predictable.

The probability distribution for amplitude $R(t)$ and phase $\phi(t)$ at a fixed time t can be computed directly from the Gaussian distribution for $x(t)$ and $y(t)$ suggested above. It is usually assumed that both these components have equal variances σ^2 in the far field scattering range of beyond-the-horizon propagation.

$$P(x,y)dxdy=\frac{dxdy}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp-\left\{\frac{x^2+y^2-2\rho xy}{2\sigma^2(1-\rho^2)}\right\},$$

where

$$\langle x^2\rangle=\langle y^2\rangle=\sigma^2$$

$$\langle xy\rangle=\sigma^2\rho.$$

Transformation to the polar coordinates of (4.1.1) yields the joint amplitude-phase distribution.

$$P(R,\phi)\,RdRd\phi=\frac{RdRd\phi}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp-\left\{\frac{R^2[1-\rho\sin^2\phi]}{2\sigma^2(1-\rho^2)}\right\} \tag{4.1.2}$$

Previous discussions of random signals have taken the cross-correlation (ρ) between the in and out-of-

phase components to be zero. This is justified for electrical shot noise which is a randomly phase-modulated signal. It is not justified for scatter signals, which are essentially amplitude-modulated. This cross correlation has been computed explicitly for line-of-sight propagation geometries [40], and is definitely not zero in that case. Although a similar calculation needs to be performed for scatter propagation geometries, it is quite possible that ρ has a finite value.

The distribution of signal amplitudes is established by integrating (4.1.2) over all values of the phase, ($-\pi$ to π).

$$P[R]RdR=\frac{dRR}{\sigma^2\sqrt{1-\rho^2}}\,I_0\left[\frac{R^2}{2\sigma^2}\,\frac{\rho}{1-\rho^2}\right]\exp-\left[\frac{R^2}{2\sigma^2(1-\rho^2)}\right] \tag{4.1.3}$$

If the cross correlation ρ were zero, this would reduce to the Rayleigh distribution for R [50, 51]. The root mean square amplitude is independent of ρ,

$$R_{rms}=\sqrt{\langle R^2\rangle}=\sigma\sqrt{2} \tag{4.1.4}$$

and forms a convenient amplitude reference.

The probability that the amplitude exceeds a given level r is obtained by integrating (4.1.3) from r to infinity, and is plotted in figure 17. The theoretical curves for different ρ are plotted on Rayleigh paper, so as to exhibit departures from the straight line plot of a simple Rayleigh distribution ($\rho=0$). This curve indicates that finite values for the cross correlation enhance large values of R relative to those of the Rayleigh distribution. The probability of observing very small signal values is correspondingly smaller than for the Rayleigh case. It is quite

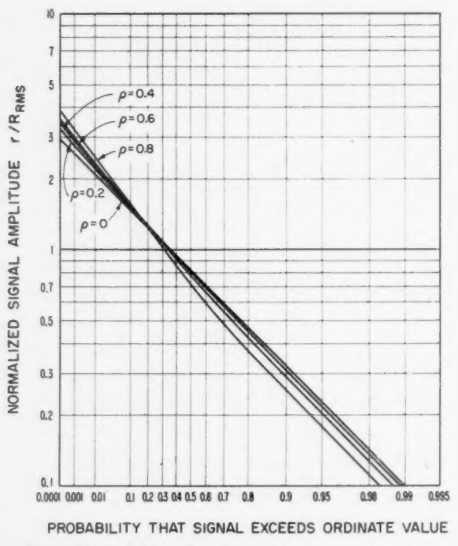


FIGURE 17. *Probability that instantaneous signal amplitude exceeds a given level measured in terms of the RMS signal level for various cross correlations ρ.*

significant that this very departure has been observed experimentally on many scatter links [47, 52]. It should also be noted, however, that a signal level trend superimposed on the random amplitude process would produce the same general behavior for finite data samples.

The possibility of a steady signal being admixed with the random scatter signal cannot be ignored, especially on short paths where the ground wave or reflecting layers may be important. The corresponding statistical problem of a constant vector plus a random process was summarized in section 3.1. It is apparent that the entire signal level distribution is raised by such an addition, which offers a third possible explanation for the non-Rayleigh character of the signal level distributions.

The average signal level can be computed from (4.1.3).

$$\langle R\rangle=\sigma\,\frac{\sqrt{\pi}}{2}\left[\sqrt{1+\rho}\,\frac{2}{\pi}\,E\left(\frac{\sqrt{2\rho}}{1+\rho}\right)\right], \tag{4.1.5}$$

where $E(x)$ is the complete elliptic function of the second kind. The correlation-dependent factor in square brackets varies by only ten percent over the entire range of ρ, so the Rayleigh form is an adequate expression.

The distribution of signal phases can be established by averaging (4.1.2) over all possible amplitudes.

$$P[\phi]d\phi=\frac{d\phi}{2\pi}\,\frac{\sqrt{1-\rho^2}}{1-\rho\sin 2\phi} \tag{4.1.6}$$

This is plotted in figure 18 for $0<\phi<\pi$ (since it is symmetric about the origin) and various ρ, indicating the departure from a uniform phase distribution which is usually associated with a Rayleigh distribution of amplitudes.

The foregoing description applies to the signal fluctuations observed during an interval ranging from several minutes to several hours. A substantial amount of experimental data is presented in terms of distributions of hourly medians, which are found to be "log-normally" distributed [53]. McCrossen [54] used this experimental result to suggest a phenomenological theory of nonstationary signal distributions for all time spans. The essential idea is to treat the signal variance in eq (4.1.3) as a long-term random variable. It is assumed that σ is sufficiently constant over intervals required for the averaging of short-term fluctuations. The quantity which is observed to be log-normally distributed is the hourly median of received power,

$$Z=P_{0.5}=\sigma^2\log 4=\langle R^2\rangle\log 2, \tag{4.1.7}$$

where ρ is taken to be zero for convenience. The distribution for Z is thus

$$P(Z)dZ=\frac{1}{\sqrt{2\pi}\zeta}\,\frac{dZ}{Z-Z_0}\exp-\left[\frac{[\log(Z-Z_0)-\mu]^2}{2\zeta^2}\right]. \tag{4.1.8}$$

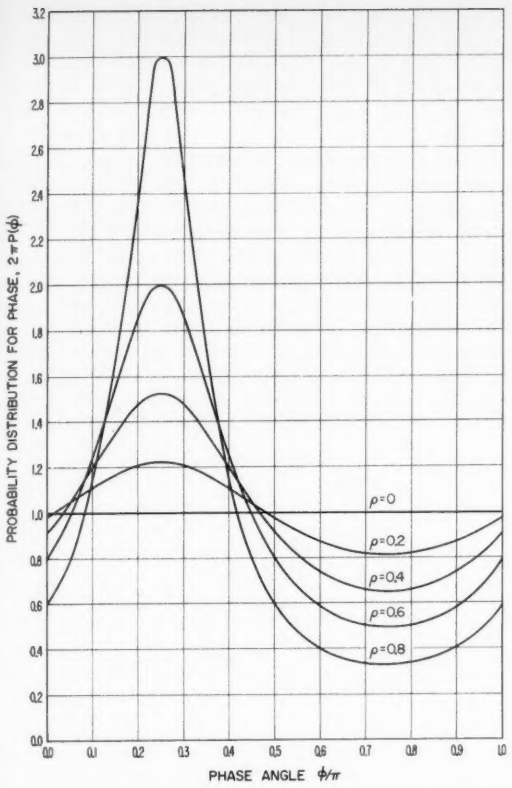


FIGURE 18. *Distribution of scatter signal phase angle ϕ for various cross correlations ρ.*

The corresponding short-term amplitude fluctuation must now be considered as a conditional probability density; that is, given σ or Z, the distribution of R is,

$$P(R|Z)=\frac{dRR}{Z}\ \log 4\ \exp-\left\{\frac{R^2\log 2}{Z}\right\}. \quad (4.1.9)$$

McCrossen [54] has evaluated the probability that R exceeds a fixed level r_0 at any instant, by integrating this expression over all possible values of Z, as well as $R>r_0$.

4.2. Electromagnetic Scattering

The basic field eq (3.2.3) for propagation through a turbulent medium also generates the solutions which describe propagation beyond the optical horizon. The single scattering (Born) approximation (3.2.4) is consistently used to describe scatter propagation [7, 45, 46]. E_0 is the field which would be received at $\vec{R}$ if there were no refractive irregularities, and is therefore the spherical earth diffraction field at points beyond the horizon. It has been established experimentally [47] that the scatter field strengths are many orders of magnitude greater than the diffracted fields 50 km beyond the horizon, so that E_0 must be omitted as the source of the scatter signals. This leaves the integral term in eq (3.2.4), which we have identified with single scattering of direct rays by the refractive blobs, as in figure 19a. The initial field E_0 is that which would be received at the scattering blob by line-of-sight propagation from the transmitter. The scattered wave proceeds from irregularity to receiver along the second (dashed) line-of-sight path, and is thus generically similar to the relay link problem characterized by figure 3a. The crucial difference lies in the scattering region V, which is now the common (wedge-shaped) volume defined by the intersection of tangent planes at the transmitter and receiver (see fig. 19a).

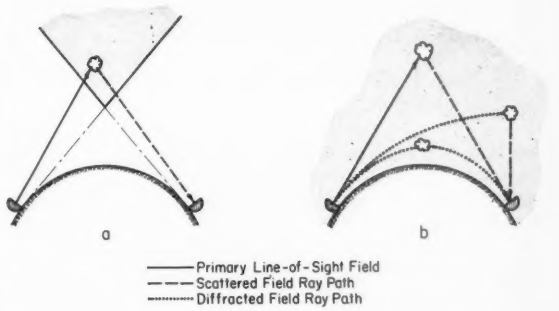


FIGURE 19. *Ray paths for scatter propagation: (a) Common (wedge) volume for line-of-sight ray paths, and (b) scattering of diffracted wave in shadow zone.*

The scattered field represented by the integral term in the Born approximation (3.2.4) gives rise to orthogonal signal components $x(t)$ and $y(t)$ which are in- and out-of-phase respectively relative to the comparatively weak phase reference E_0.

$$x+iy=FE_s=Fk_f^2\int_V d^3rG(\vec{R},\vec{r})\Delta\epsilon(\vec{r},t)E_0(\vec{r})\ \sin\chi, \quad (4.2.1)$$

F is the receiver gain factor,

$$F^2=A_R=\frac{\lambda^2G_R}{4\pi}$$

which transforms field strength (volts/cm) into receiver voltage (volts). We omit the polarization factor sin χ for the time being. Separating (4.2.1) into real and imaginary components

$$x(t)=F\int_V d^3r\Delta\epsilon(\vec{r},t)\ \mathrm{Re}\left[-k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right],$$

$$y(t)=F\int_V d^3r\Delta\epsilon(\vec{r},t)\ \mathrm{Im}\left[-k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right], \quad (4.2.2)$$

completes the identification of voltage components in figure 16a and eq (4.1.1).

The variance and cross correlation of the orthogonal voltage components required in (4.1.2) can be established from these expressions using the spectrum method, as given by (2.3.1). Interchanging the orders of wavenumber (k) and scattering volume (r) integrations yields compact results similar to (3.4.1) in which the turbulence model $S(k)$ and geometrical considerations are explicitly separated.

$$\langle x^2\rangle=\frac{F^2}{8\pi^3}\int d^3kS(k)\left|\int_V d^3re^{i\vec{k}\cdot\vec{r}}\,\mathrm{Re}\left[k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right]\right|^2$$

$$\langle y^2\rangle=\frac{F^2}{8\pi^3}\int d^3kS(k)\left|\int_V d^3re^{i\vec{k}\cdot\vec{r}}\,\mathrm{Im}\left[k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right]\right|^2$$

$$\langle xy\rangle=\frac{F^2}{8\pi^3}\int d^3kS(k)\left\{\int_V d^3re^{i\vec{k}\cdot\vec{r}}\,\mathrm{Re}\left[k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right]\right\}$$

$$\times\left\{\int_V d^3re^{i\vec{k}\cdot\vec{r}}\,\mathrm{Im}\left[k_f^2G(\vec{R},\vec{r})E_0(\vec{r})\right]\right\} \quad (4.2.3)$$

The last integral is real in virtue of the Hermitian property of the spectrum $S^*(k)=S(-k)$, which follows from eq (2.3.1). The two basic scatter propagation integrals contained in (4.2.3) have not been evaluated explicitly to date. They have been attacked by the author at two levels of sophistication. The first method employs the simple common volume cutoff indicated by figure 19a. $E_0(r)$ is taken as a spherical wave expanding from the transmitter, and $G(R,r)$ is the free space Green's function of eq (3.2.5), representing unshielded expansion of the scattered wave from the blob to receiver. The common volume is then the wedge-shaped region described above and it is convenient to use rectangular coordinates centered on the wedge apex and the great circle plane between transmitter and receiver.

The second approach is indicated by figure 19b and recognizes the role of both line-of-sight and diffracted primary and/or scattered waves. In this case, one takes E_0 to be the actual (series) solution for a dipole radiating above a spherical conducting earth. This automatically includes line-of-sight propagation to blobs in the common volume, as well as blobs below the horizon which are reached only by the diffracted component of E_0. Subsequent propagation of the scattered radiation must also recognize the earth-screening effect, which is to say that the Green's function $G(R,r)$ must be that appropriate to the exterior of a conducting sphere. However, this is just the field due to a unit dipole placed at the scattering point, and is therefore given again by a series solution of the form used for E_0. This approach treats the three types of scatter paths shown in figure 19b exactly, and includes the scattering of diffracted waves and diffraction of scattered diffracted waves in a natural fashion. This refinement is probably of academic interest only, since Arons [55] has concluded that such effects are quite small for normal scatter circuits. This approach does have the advantage of treating height-gain effects and near horizon paths in a unified manner. The volume integrations in (4.2.3) required for this approach are not trivial and the surviving summations must be performed by the Watson transformation [56] in the end.

All of the precise development described above is essentially new and unpublished. The considerable theoretical literature on scatter propagation is based on the cross section approach which is built on a series of approximations to the integral representation (4.2.1) for the (complex) scattered field, and treats only the average scattered power $\langle R^2\rangle=\langle x^2\rangle+\langle y^2\rangle$. The first step is the far field approximation, which notes that the scattering blobs are always many, many wavelengths from both the transmitter and receiver (see, however, ref. 24). This allows the scalar distance between the receiver $\vec{R}$ and variable scattering point $\vec{r}$ to be expanded as follows:

$$\left|\overrightarrow{R_2-r}\right|\cong R_2-\frac{\vec{R}_2\cdot\vec{r}}{R_2}=R_2-\frac{\vec{k}_2\cdot r}{k_f}, \quad (4.2.4)$$

where $\vec{k}_2$ is the propagation vector of the scattered wave ($|\vec{k}_2|=k_f=2\pi/\lambda$). The field E_0 incident on the point r is representable as,

$$E_0=\sqrt{\frac{P_TG_T}{4\pi}}\;\frac{1}{R_1}e^{i\vec{k}_1\cdot\vec{r}},$$

where $\vec{k}_1$ is the upgoing (line-of-sight) propagation vector. G_T is the gain of the transmitter along $\vec{k}_1$, and P_T is the transmitted power. Combining these expansions with (4.2.1) yields:

$$E_s\cong\frac{\pi}{\lambda^2}\sqrt{P_T}\,e^{ik_fR_2}\int_V d^3r\,\frac{e^{i\vec{r}\cdot(\overrightarrow{k_1-k_2})}}{R_1R_2}\Delta\epsilon(\vec{r},t)\sqrt{\frac{G_T}{4\pi}} \quad (4.2.5)$$

The mean square received power is given in terms of receiver aperture $A_R=G_R\lambda^2/4\pi$ by

$$P_R=A_R\langle|E_s|^2\rangle=\frac{P_T}{16\lambda^2}\int d^3r\int d^3r'\langle\Delta\epsilon(\vec{r},t)\Delta\epsilon(\vec{r}',t)\rangle$$

$$\sqrt{G_RG_R'G_TG_T'}\;\frac{e^{i\vec{r}\cdot(\overrightarrow{k_1-k_2})}e^{-i\vec{r}\cdot(\overrightarrow{k_1'-k_2'})}}{R_1R_2R_1'R_2'} \quad (4.2.6)$$

where k_1, k_2 and the gain functions are implicit functions of position. The second approximation is now invoked by noting that the space correlation in (4.2.6) vanishes unless the points $\vec{r}$ and $\vec{r}'$ are within several hundred meters of one another. The other factors in the integrand are relatively insensitive functions of position, so that

$$R_1=R_1',\ R_2=R_2',\ k_1-k_2=k_1'-k_2',\ G_T=G_T',\ G_R=G_R'$$

over the important joint regions of integration. By transforming to sum and difference coordinates,

$$r=r_1+r_2 \text{ and } \rho=r_1-r_2,$$

one can rewrite (4.2.6) as

$$P_R=\frac{P_T}{16\lambda^2}\int_V d^3r\,\frac{G_R G_T}{R_1^2R_2^2}\int_V d^3\rho e^{i\vec{\rho}\cdot(\vec{k_1}-\vec{k_2})}\langle\Delta\epsilon^2\rangle C(\rho).$$

Because $C(\rho)$ is nonvanishing only over a small volume of ρ space, the ρ integration can be extended over all space without misrepresentation. The Fourier inverse of (2.3.1) then indicates that this (difference) integral is nothing more than the spectrum $S(k)$ evaluated at the uniquely important "scattering wavenumber",

$$K=|\overrightarrow{k_1-k_2}|=\frac{4\pi}{\lambda}\sin\theta/2, \tag{4.2.7}$$

where θ is the scattering angle formed by line-of-sight rays to the volume element d^3r,

$$P_R=\frac{P_T}{16\lambda^2}\int_V d^3r\,\frac{S\left[\left(\frac{4\pi}{\lambda}\sin\frac{\theta_r}{2}\right)\right]}{R_1^2R_2^2}G_TG_R. \tag{4.2.8}$$

This expression indicates that the scattering process selects those blobs in the hierarchy described by $S(k)$ whose wave numbers satisfy (4.2.7). This corresponds to constructive interference of the waves diffracted by the appropriate grating (i.e., blob) spacing for the Bragg angle θ. This selective isolation of only a certain size class of blobs is equivalent to the narrow band filter shown in figure 20 applied to the spectrum of irregularities.

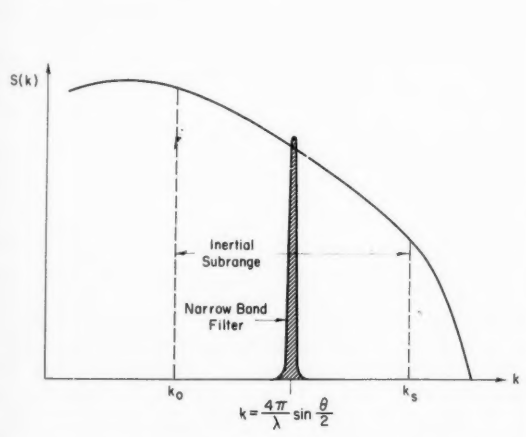


FIGURE 20. *Typical spectrum of refractive irregularities showing narrow band filter at the scattering wavenumber which characterizes tropospheric scattering.*

θ is the average scattering angle and λ is the radiation wave length.

The expression (4.2.8) for received power is often quoted in terms of a scattering cross section

$$\sigma(\theta,\lambda)=\frac{\pi^2}{\lambda^4}S\left(\frac{4\pi}{\lambda}\sin\frac{\theta}{2}\right). \tag{4.2.9}$$

This cross section measures the energy scattered per unit volume V, into a unit solid angle $d\omega$ which makes the angle θ with the initial Poyting vector S_0. This is illustrated in figure 21. The scattering cross section is always defined for a unit incident power, so that at a distance R from the scattering element V,

$$\sigma=\frac{R^2}{V}\left\langle\left|\frac{E_s(R)}{E_0}\right|^2\right\rangle, \tag{4.2.10}$$

which is equivalent to (4.2.9) in virtue of the Born approximation (4.2.1) for E_s. σ has the units cm^{-1}, since it is an ordinary cross section (cm^2) per unit volume (cm^{-3}).

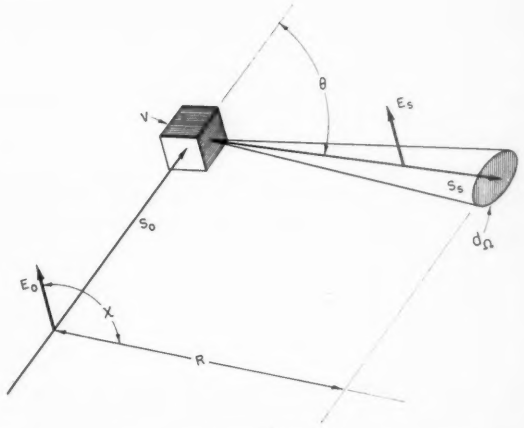


FIGURE 21. *Geometrical description of scattering cross section σ per unit volume and solid angle $d\Omega$ per unit incident power.*

The angular dependence of σ is often interpreted as the scattering polar diagram of an average blob. Reference to table 1 shows that almost all of the scattered radiation is projected into a small cone of beamwidth λ/l_0 about the forward direction. This means that scatter propagation circuits depend on the extremely weak side lobes of the scattering blobs, which explains, in part, why large dishes and high power are required to operate such links. Feshbach [57] has suggested that many small angle multiple scatterings could be more efficient than a single large angle scattering. Such effects certainly become increasingly important at frequencies above 1,000 Mc and/or very long paths, since the convergence of the Born series is related to the smallness of the equivalent line-of-sight rms phase variation computed along the equivalent (kinked) ray path eq (3.2.11) with $L\simeq d$. If this view is correct, the entire subject, which is summarized in subsequent

sections, requires revision. Further research should be directed to this interesting point.

The equivalence between σ and S given by eq (4.2.9) can be used to produce the familiar expression for received power $P_R \sim \langle R^2 \rangle$ to transmitted power $P_T \sim |E_0|^2$

$$\frac{P_R}{P_T} = \frac{\lambda^2}{16\pi^2} \int_V d^3r \, \frac{\sigma(\theta,\lambda)}{R_1^2 R_2^2} \, G_T G_R. \qquad (4.2.11)$$

The corresponding spectrum integral is more fundamental, since the joint frequency-distance dependence is emphasized thereby.

$$\frac{P_R}{P_T} = \frac{1}{16\lambda^2} \int_V d^3r \, \frac{G_T G_R}{R_1^2 R_2^2} \, S\left(\frac{4\pi}{\lambda} \sin\frac{\theta}{2}\right) \qquad (4.2.12)$$

Data taken with very narrow beams (see fig. 16b) is usually expressed in terms of this ratio.[6] If the beams are so narrow that the scattering angle, and hence the spectrum, is substantially constant over the common volume, the integration can be collapsed to

$$\frac{P_R}{P_T} \simeq \frac{V G_T G_R}{\lambda^2 d^4} \, S\left(\frac{4\pi}{\lambda} \sin\frac{\theta}{2}\right), \qquad (4.2.13)$$

since $R_1 \simeq R_2 \simeq d/2$. The scattering volume is set by the aerial beamwidths β and the transmission distance d by $V \simeq 2(\beta d/4)^3/\theta$, where the average scattering angle is $\theta = d/a$ for $d \ll a$.

$$\frac{P_R}{P_T} \simeq \frac{\beta^3 G_T G_R}{32\lambda^2 d^2} \, S\left(\frac{2\pi d}{\lambda a}\right). \qquad (4.2.14)$$

The majority of broad-beam data is given with respect to the free-space power P_{FS} which would be received over a line-of-sight path of the same length d.

$$\frac{P_{FS}}{P_T} = \frac{G_T}{4\pi d^2} \, \frac{\lambda^2 G_R}{4\pi}. \qquad (4.2.15)$$

This measure of transmission efficiency emphasizes the loss which is due to the scattering propagation agency alone,

$$\frac{P_R}{P_{FS}} = \pi^2 d^2 \int_V d^3r \, \frac{1}{R_1^2 R_2^2} \, S\left(\frac{4\pi}{\lambda} \sin\frac{\theta}{2}\right). \qquad (\quad .16)$$

The receiver and transmitter gains also cancel in this expression, since they are relatively constant over the scattering volume. The equivalent expressions (4.2.12) and (4.2.16) will be used interchangeably in what follows.

[6] Additional factors of two are sometimes included [49] in such expressions to represent the addition of ground reflected waves. Our gain factors are defined to include all such effects.

4.3. Radiofrequency Dependence

The variation of received power with carrier frequency can be established without evaluating the difficult volume integrals presented in the last section. This is because the value of λ and θ relevant to most scatter paths combine to emphasize "scattering-wave numbers" which lie in the inertial range of the spectrum of tropospheric irregularities.[7] The physical models for $S(k)$ are proportional to simple inverse powers of k in the inertial range, and the radiofrequency dependence of the power ratio expressions (4.2.12) or (4.2.16) can be brought outside the integrals. If $S(k) = S_0 k^{-n}$, one has from (4.2.12),

$$\frac{P_R}{P_T} = \frac{\lambda^{n-2}}{16} \int_V d^3r \, \frac{G_T G_R}{R_1^2 R_2^2} \, \frac{S_0}{\left(4\pi \sin\frac{\theta}{2}\right)^n} \qquad (4.3.1)$$

There is still a frequency dependence concealed in the gain functions, which is due to both wave length-dependent height gain factors and the absorbing area-gain relationship $G_R = A_R \, 4\pi/\lambda^2$. The ratio of received powers for scaled aerials on the same path is therefore

$$\frac{P_R(\lambda_1)}{P_R(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^{n-4} \qquad (4.3.2)$$

The mixing-in-gradient model (2.4.6) corresponds to $n = 5$ and predicts a linear-wavelength dependence of scattered power. This prediction is in good agreement with careful analysis of the data from numerous broad-beam experiments when the role of height gain factors is carefully included [58]. The same model implies a λ^3 variation for the ratio P_R/P_T, which is quite consistent with narrow beam data [59]. It would appear that the $\lambda^{-0.33}$ dependence predicted by the Obukhov model (2.4.4) is not consistent with present radio data.

Bean [60] has studied the effect of tropospheric absorption on scatter signals as a function of frequency and distance, and finds that such effects do not influence the above conclusions. The variability of the frequency scaling exponent has been noted on all links, and one must be quite explicit about just which quantities are being compared on the two frequencies. It would seem that ratios of hourly means on the same path for the same hours would be the best choice of data, since [1] two widely separated frequencies do not fade together (see section 4.9), and [2] diurnal and seasonal effects may affect two frequencies differently. On the other hand, the NBS group finds that hourly medians on adjacent identical paths are not always the same, and a further examination of the problem is indicated. The above statements about frequency dependence refer to the average power, and hence to the average frequency exponent. It should be emphasized that it is this average value, and not the extreme values, [58, 61] which should be correlated with present (smoothed) experimental data.

[7] For example, if $f = 1{,}000$ Mc and $d = 400$ miles, $\lambda = 1$ ft and $\theta = 0.1$ radian eq (4.2.7) gives $K = 2$ m^{-1}.

4.4. Distance Dependence

A large amount of experimental data on signal-power attenuation relative to free space drawn from reference [62] is plotted in figure 22. This is a heterogeneous sample of long-term medians, obtained in most cases with broad beam antennas. The data were taken during the wintertime, mostly in the afternoon, when tropospheric scatter is expected to be the dominant propagation mode. These data have not been normalized for the effects of radiofrequency, antenna heights, terrain or meteorological influences, which explains the considerable data scatter. If the observed linear wave-length dependence of the scatter signals is used to reduce these data to a common radiofrequency base, the data scatter is reduced. The empirical curve drawn through the data corresponds to an exponential variation of signal power at large distances, with approximately 0.1 db/mile. At shorter distances, the distance dependence is more complicated.

The distance dependence of scattered power is not as simple to extract as the radiofrequency variation. This is because such predictions depend upon careful integration of the spectrum's angular dependence over the common volume V, which, in turn, is defined by the transmission distance d. The gain factors are essentially constant in (4.2.12), so that the distance dependence of the mixing-in-gradient model (2.2.4) can be expressed in terms of the integral

$$P_R \simeq \int_V d^3r \, \frac{1}{R_1^2 R_2^2} \, \frac{\left(\frac{d\epsilon_0}{dh}\right)^2}{\left(\sin\frac{\theta}{2}\right)^5}, \qquad (4.4.1)$$

where all factors in the integrand depend on the great circle path distance d. The mixing-in-gradient model is subject to a more severe test than other models because it contains no (adjustable) turbulence parameters. The NBS group has used the meteorologically observed exponential variation of $d\epsilon_0/dh$ up to stratospheric heights to predict the experimental data presented in figure 22 quite well out to 700 miles [61, 62].

A substantial amount of early theoretical prediction [58, 63, 64] was based on height dependences assumed for the mean square dielectric fluctuation $\langle \Delta\epsilon^2 \rangle \propto h^{-n}$. This also increased the distance scaling law exponent by n and supplied some of the missing exponents for older correlation function models thereby. Insofar as one believes in the physical theories of mixing models, it is more realistic to insert measured profiles for $\epsilon_0(h)$ and let the theories be tested thereby.

In making precise comparisons between theory and experiment, it is necessary to recognize several important details. The effect of (irregular) terrain obstacles in modifying the takeoff angles, and hence the minimum scattering angle versus distance relation, led the NBS group to the concept of effective angular distance [58]. Refraction of the upgoing and downgoing beams by the mean profile ϵ_0 has been analyzed thus far by the equivalent earth radius method [58]. Bean [60] established theoretically the corrections due to tropospheric absorption on distance dependent scaling laws.

The simplest case to analyze is that for narrow-antenna beams, since the volume integration is then collapsed and approximated by eq (4.2.14). The

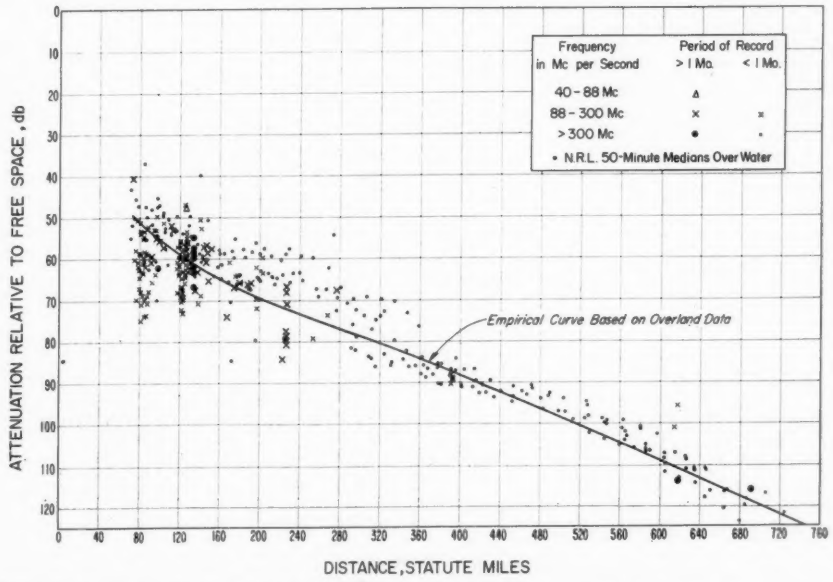


FIGURE 22. *Experimental data on distance dependence of tropospheric scatter field strength relative to free space for all frequencies, after Rice, Longley, and Norton* [62].

mixing-in-gradient model (2.4.6) predicts the following distance scaling law,

$$\frac{P_R}{P_T} \simeq \left[\frac{d\epsilon_0}{dh}\right]_V^2 \frac{\text{constant}}{d^7}, \qquad (4.4.1)$$

for identical antennas and height gain effects. This result is independent of turbulent parameters, whereas the Obukhov model (2.4.4) gives

$$\frac{P_R}{P_T} \simeq \left[\frac{d\epsilon_0}{dh}\right]_V^2 \frac{k_0^{2/3}(V)}{d^{5.7}} \text{ constant}. \qquad (4.4.2)$$

Both models depend upon the gradient of the mean profile ϵ_0 at the average scattering altitude. This point rises as the path length is increased and the gradient probably decreases accordingly. The narrow-beam data [65] suggests a scaling law of the form $1/d^9$, or possibly even $e^{-\gamma d}$. Appropriate fitting of $d\epsilon_0/dh$ and/or k_0 in the foregoing can bring either theoretical result into agreement with almost any experimental scaling law, so that this type of measurement cannot be considered a very severe test of the theory at the present time.

4.5. Absolute Power Estimates

Probably the most difficult problem in scatter theory is to make reliable absolute power estimates. All of the explicit and implicit factors in (4.2.12) must be carefully evaluated. The following points must be recognized in reaching accurate power predictions: (a) Horizon limitations of local terrain, (b) height gain effects in the antennas, (c) lobe patterns of the antennas, (d) variation of dielectric irregularity intensity with height, (e) refraction of primary and scattered rays, (f) absorption, (g) anisotropy influences, and (h) transmitter and receiver antenna loading.

The problem has only been solved completely by the NBS group, using numerical techniques [61, 62]. They considered broad-antenna patterns with the mixing-in-gradient model and obtained good (absolute) agreement with experimental data.

The problem is primarily one of finding a judicious coordinate system in which to perform the complicated volume integration of (4.2.12). The NBS spherical system [66] can be integrated exactly for many turbulence models [67], but is not adaptable to height-dependent scattering cross sections. The rectangular system used by Staras [64] circumvents this problem, but relies on approximations which are difficult to assess. It is anticipated that the precise integration schemes suggested by eq (4.2.3) will supply an accurate means for making absolute power level comparisons. A height-dependent factor $(d\epsilon_0/dh)^2$ in the spectrum enters into (4.2.3) as a factor $d\epsilon_0/dh$ in each fundamental integral.

An important achievement for scatter theory is the reliable prediction of signal strengths beyond the horizon from surface measurements of $(d\epsilon_0/dh)_0$ at the earth's surface [62]. This means that diurnal and seasonal (relative) variations of scattered power can be anticipated without recourse to complicated volume integrations, since the gradient profile,

$$\frac{d\epsilon_0}{dh} = \left[\frac{d\epsilon_0}{dh}\right]_0 F(h), \qquad (4.5.1)$$

permits the time-variable ground level factor to be removed from the geometrical integrals.

4.6. Amplitude Stability of Scatter Signals

The characteristic random fading of the amplitude of scatter signals is of considerable practical importance, since useful information can only be transmitted when the signal strength is above a fixed level. The temporal variations of the signal envelope are physically related to the incoherent addition of (random) Doppler-shifted waves received from different elements of the scattering volume. A typical time record of scatter signals is shown in figure 23, together with the time-displaced vector voltage diagrams corresponding to the instants t_1 and t_2. The basic theoretical problem is to predict the joint probability that the signal amplitude will assume the values R_1 and R_2 at times separated by an interval $\tau = t_2 - t_1$. The average time between maxima and zero crossings, amplitude auto correlations, etc., can all be computed from such a result.

The purely statistical descriptions of fading must be discussed first. The distribution of time-displaced signal amplitudes shown in figure 23 depends on the four dimensional probability density for the set (x_1, y_1, x_2, y_2). Since the individual orthogonal signal components $x(t)$ and $y(t)$ are considered to be distributed in a Gaussian manner at each instant, the required density is a four dimensional Gaussian multivariate form,

$$P(Z_i) = \frac{1}{4\pi^2 |M|} \exp -\frac{1}{2} \sum_1^4 \sum_1^4 Z_i Z_j M_{ij}^{-1} \qquad (4.6.1)$$

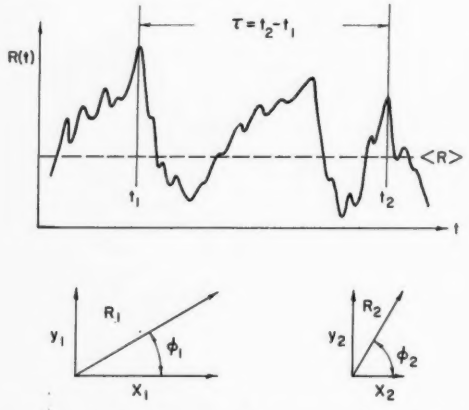


FIGURE 23. *Typical time record of randomly fading scatter signal amplitude with corresponding time-displaced vector voltage diagrams.*

(R) is the average amplitude and ϕ is the instantaneous phase of the signal.

where $Z_1=x(t)$, $Z_2=y(t)$, $Z_3=x(t+\tau)$, $Z_4=y(t+\tau)$. The moment matrix M_{ij} is the basic characterization [8] of the fading signal.

$$M=\begin{bmatrix} \langle x^2(t)\rangle & \langle x(t)y(t)\rangle & \langle x(t)x(t+\tau)\rangle & \langle x(t)y(t+\tau)\rangle \\ \langle x(t)y(t)\rangle & \langle y^2(t)\rangle & \langle x(t+\tau)y(t)\rangle & \langle y(t)y(t+\tau)\rangle \\ \langle x(t+\tau)x(t)\rangle & \langle x(t+\tau)y(t)\rangle & \langle x^2(t+\tau)\rangle & \langle x(t+\tau)y(t+\tau)\rangle \\ \langle x(t)y(t+\tau)\rangle & \langle y(t)y(t+\tau)\rangle & \langle x(t+\tau)y(t+\tau)\rangle & \langle y^2(t+\tau)\rangle \end{bmatrix} \qquad 4.6.2$$

The matrix M_{ij} is symmetrical because the component processes are assumed to be stationary. The cylindrical transformations $x=R\cos\phi$ and $y=R\sin\phi$ which define the signal amplitude and phases (see fig. 23) must be inserted into expression (4.6.1) to establish the time-displaced phase and amplitude distribution. This portion of the problem is analytically very complicated unless the moment matrix M_{ij} assumes especially simple forms, and it is desirable to investigate the time correlations themselves before pursuing the statistical problem further.

The propagation problem is to compute the moment matrix of eq (4.6.2) explicitly. This has not been done. One can, however, establish general expressions for the time-displaced correlations [9] in M_{ij} by using the integral representations for $x(t)$ and $y(t)$ established previously from the Born approximation (4.2.2), viz,

$$x(t)=\int_V d^3r\Delta\epsilon(r,t)H(r),$$

$$y(t)=\int_V d^3r\Delta\epsilon(r,t)J(r), \qquad (4.6.3)$$

where $H(r)$ and $J(r)$ are the real and imaginary parts respectively of the complex product $[k^2G(R,r)E_0(r)]$ in eq (4.2.2). The appropriate averages can now be computed using the integral representation (2.5.2) for the space- and time-displaced correlation of dielectric fluctuations.

$$\langle x(t)x(t+\tau)\rangle=\int_V d^3rH(r)\int_V d^3r'\,H(r')\langle\Delta\epsilon(r,t)\Delta\epsilon(r',t+\tau)\rangle$$

$$=\frac{1}{(2\pi)^3}\int d^3k\,S(k)\eta\,[k,\tau]e^{-i\vec{k}\cdot\vec{U}\tau}\left|\int_V d^3rH(r)e^{i\vec{k}\cdot\vec{r}}\right|^2$$

$$\langle y(t)y(t+\tau)\rangle=\frac{1}{(2\pi)^3}\int d^3k\,S(k)\eta(k,t)\,e^{-i\vec{k}\cdot\vec{U}\tau}\left|\int_V d^3rJ(r)\,e^{i\vec{k}\cdot\vec{r}}\right|^2$$

$$\langle x(t)y(t+\tau)\rangle=\langle x(t+\tau)y(t)\rangle$$

$$=\frac{1}{(2\pi)^3}\int d^3k\,S(k)\eta(k,t)\,e^{-i\vec{k}\cdot\vec{U}\tau}\left(\int_V d^3rH(r)e^{i\vec{k}\cdot\vec{r}}\right)\left(\int_V d^3r'J(r')e^{-i\vec{k}\cdot\vec{r}'}\right) \qquad (4.6.4)$$

The important point to note here is that the geometrical integrals involving $H(r)$ and $J(r)$ are precisely those required for the same-time calculations of signal properties. Although these integrals have not been evaluated to date, it is significant that the two-time problem is no more difficult than the same-time problem as far as the propagation calculations are concerned.

The statistical problem has been treated heretofore [51, 69] by assuming very special forms for the moment matrix.

$$\langle x_1^2\rangle=\langle x_2^2\rangle=\langle y_1^2\rangle=\langle y_2^2\rangle=\sigma^2$$
$$\langle x_1x_2\rangle=\langle y_1y_2\rangle=\sigma^2\mu(\tau) \qquad (4.6.5)$$

$\mu(\tau)$ is here called the basic propagation correlation coefficient. All cross correlations are assumed to vanish identically. The joint probability density (4.6.1) now reduces to a comparatively simple form, which is readily transformed into the polar amplitude-phase coordinates of figure 23,

$$P[R_1,R_2,\phi_1,\phi_1;\tau]=\frac{R_1R_2}{4\pi^2\sigma^4(1-\mu^2)}$$
$$\exp-\left[\frac{R_1^2+R_2^2-2\mu R_1R_2\cos(\phi_1-\phi_2)}{2\sigma^2(1-\mu^2)}\right] \qquad (4.6.6)$$

The joint amplitude distribution for R_1 and R_2 is obtained by integrating this expression over both phase angles.

$$P[R_1,R_2;\tau]=\frac{R_1R_2}{\sigma^4(1-\mu^2)}\,I_0\left[\frac{\mu R_1R_2}{\sigma^2(1-\mu^2)}\right]e^{\frac{R_1^2+R_2^2}{2\sigma^2(1-\mu^2)}} \qquad (4.6.7)$$

[8] $|M|$ is the determinant of M_{ij} and M_{ij}^{-1} its inverse.
[9] The one-time components are special cases of the time-displaced correlations, viz., $\langle x(t)y(t)\rangle=\langle x(t)y(t+\tau)\rangle_{\tau=0}$.

The amplitude autocorrelation can be expressed in terms of complete elliptic functions,

$$\langle R(t)R(t+\tau)\rangle=\langle R_1R_2\rangle=\sigma^2\{2E(\mu)-(1-\mu^2)K(\mu)\} \tag{4.6.8}$$

and is plotted in figure 24 as a function of $\mu(\tau)$. For large time separations (i.e., μ small), this result approaches the square of the average amplitude, $\langle R\rangle^2=\frac{\pi}{2}\sigma^2$. For small separations or μ near unity, it approaches $\langle R^2\rangle=2\sigma^2$. The frequency spectrum of amplitude variations can be obtained from (4.6.8) by the Fourier cosine transformation, if $\mu(\tau)$ is known explicitly as a result of propagation calculations.

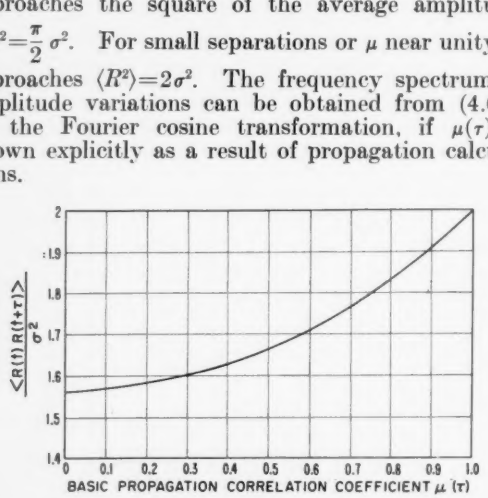


FIGURE 24. *Normalized autocorrelation of scatter signal amplitude versus basic propagation correlation coefficient* $\mu(\tau)$.

The envelope fading rate, or average number of positive crossings of the median signal level R per unit time in figure 23, is a convenient measure of the time-varying random signal. This fading rate N can be identified in the small argument expansion of the amplitude autocorrelation [30].

$$\langle R(t)R(t+\tau)\rangle=\langle R^2\rangle[1-\tfrac{1}{2}(\pi N\tau)^2\ \ldots\ ]. \tag{4.6.9}$$

If the basic propagation correlation coefficient $\mu(\tau)$ has a similar expansion,

$$\mu(\tau)=1-\tfrac{1}{2}[\gamma\tau]^2\ \ldots \tag{4.6.10}$$

Reduction of expression (4.6.8) leads to the following relationship:

$$N=\frac{\gamma}{\sqrt{2}} \tag{4.6.11}$$

The basic propagation problem is to predict the radiofrequency and distance dependence of γ.

Early studies of fading rates on scatter links recognized that the primary fading mechanism was due to Doppler-shifting of the signal by motion of the scattering medium. An rms sum of the drift velocity U and average eddy velocity V_0 was imagined to give the effective speed [32, 68]

$$N=\frac{f}{c}\left[\left(2U\sin\frac{\theta}{2}\right)^2+\frac{1}{3}V_0^2\right]^{\frac{1}{2}}, \tag{4.6.12}$$

where θ is the average scattering angle in the common volume. The linear frequency dependence is characteristic of simple Doppler-shifting.

The geometry of a broad-beam scatter link shown in figure 25 indicates, however, that the situation is rather more complicated [69]. Consider first the midpoint (A) in the scattering volume. Since Doppler shifts are only caused by net elongation (or contraction) of the propagation path, the predominantly horizontal drift will cause no fading of the signal scattered from midpoints. On the other hand, the smaller random velocity V_0 has an arbitrary direction and can impose Doppler shifts on this component of the scattered signal. Consider next an offcenter scattering element (B). The drift velocity U now has a small component $U_\perp$ which bisects the propagation paths, and therefore does impose a Doppler shift. The fading contribution of these off-axis blobs is small because: (*i*) $U_\perp$ is a small projection of U, and (*ii*) the power scattered from point B is considerably smaller than that from point A due to the larger scattering angle θ'. Direct confirmation of this prediction was obtained by the Lincoln Laboratory Group [59], who noted that the fading rate decreased when very narrow-antenna beams are swung simultaneously away from the path midpoint A.

A considerable advance in analyzing this fading phenomenon was made recently by using the spectral representation of the turbulent dielectirc fluctuations [14]. The explicit separation of drift and random velocity contribution to the space-time correlation of $\Delta\epsilon$ in eq (2.5.2) suggests the following time-displaced generalization of the cross section result (4.2.12),

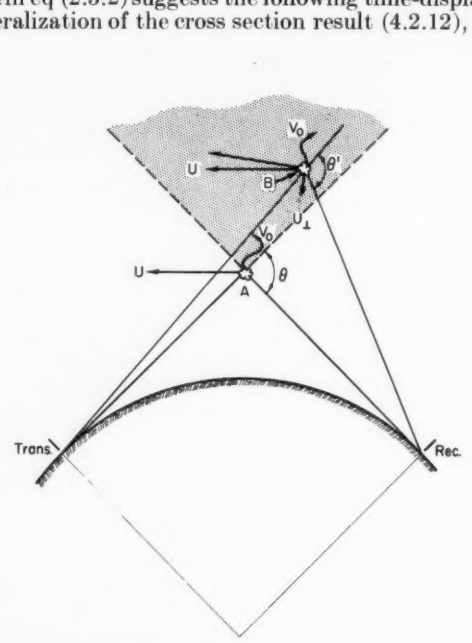


FIGURE 25. *Geometrical description of doppler shifts imposed by horizontal drift speed* $\vec{U}$ *and random velocity* $\vec{V}_0$ *at: (A) Path midpoint, and (B) off axis scattering volume elements.*

$$\langle R(t)R(t+\tau)\rangle \simeq \int_V d^3rS(|\vec{K}|)\, e^{i\vec{K}\cdot\vec{U}\tau\eta}\,[V_0k_0^{1/3}|\vec{K}|^{2/3}\tau], \tag{4.6.13}$$

where $\vec{K}$ is the scattering difference vector defined by (4.2.7). Since the major component of the drift speed is usually horizontal, and therefore perpendicular to $\vec{K}$ for midpoint scattering, the significant contribution of drift is made by volume elements far from the midpoint. The random self motion V_0 has no such projection factor and contributes throughout the scattering volume. Expansion of (4.6.13) in a power series of τ and identification with expression (4.6.9) indicates that the fading rate due to self-motion is proportional to $f^{2/3}$, while drift motion gives the usual linear Doppler dependence on frequency. Available data indicates a variable frequency exponent less than one [32], so that some mixture of the two effects is indicated. Further experimental results are needed, especially data on the distance-dependence of N. The theory needs to be clarified by evaluating $\mu(\tau)$ from eq (4.6.4) for several propagation models.

4.7. Phase Stability of Scatter Signals

The phase stability of scatter signals has received little attention to date—either theoretically or experimentally. The question is potentially quite interesting because of the possibility of passing stable frequency references beyond line-of-sight by this mode. Frequency excursions about a single carrier are completely equivalent to time rates of change of the signal phase $\phi(t)$ defined in figure 23. The mean frequency change is thus given in terms of the autocorrelation of phase

$$\langle \delta f^2\rangle = \langle \dot{\phi}(t)\dot{\phi}(t+\tau)\rangle\Big|_{\tau=0} = -\frac{d^2}{d\tau^2}\langle \phi(t)\phi(t+\tau)\rangle\Big|_{\tau=0}. \tag{4.7.1}$$

The relative change in phase between two instants,

$$\langle |\phi(t)-\phi(t+\tau)|^2\rangle \simeq 2[\langle\phi^2\rangle - \langle\phi(t)\phi(t+\tau)\rangle], \tag{4.7.2}$$

is independent of the absolute path phase reference, and is probably easier to measure experimentally.

The central quantity in such studies is the autocorrelation of phase taken between two instants. For the special assumptions indicated by (4.6.5), it is possible to derive this result in terms of the basic propagation correlation coefficient $\mu(\tau)$. Integrating (4.6.6) over the amplitudes R_1 and R_2 gives the distribution for ϕ_1 and ϕ_2 deduced by Bunimovich [70]

$$P[\phi_1,\phi_2;\tau] = \frac{(1-\mu^2)}{4\pi^2}\left[\frac{1}{1-\omega^2} + \omega\,\frac{\frac{\pi}{2}+\sin^{-1}\omega}{(1-\omega^2)^{3/2}}\right] \tag{4.7.3}$$

where $\omega=\mu\cos(\phi_1-\phi_2)$. Wagner [10] has used this result to calculate the phase autocorrelation,

$$\langle\phi(t)\phi(t+\tau)\rangle = \frac{\pi}{2}\sin^{-1}\mu + \frac{1}{2}[\sin^{-1}\mu]^2 - \pi\int_0^{\pi/2} dx[\sin^{-1}(\mu\cos x)]^2 \tag{4.7.4}$$

and this is plotted in figure 26 versus $\mu(\tau)$. For very large τ or μ small, the phases are completely uncorrelated. For μ unity, expression (4.7.4) approaches the mean square value $\langle\phi^2\rangle=\pi^2/3$.

It is interesting that the phase correlation can be inferred from experimental data on amplitude correlations, since μ is only a parameter which may be eliminated between eq (4.6.7) and (4.7.4). When the more general moment matrix (4.6.2) is used, however both results will depend on two or more basic functions of τ. In that case, it is still possible to convert phase into amplitude data and vice versa by eliminating τ, although one must first complete explicit progagation calculations for $\mu(\tau)$, etc.

4.8. Space Correlation and Diversity Reception

A combination of two or more spaced receivers is frequently exploited to overcome signal loss during deep fades. This diversity reception depends upon spacing the antennas sufficiently far apart to insure independent fading of the two signals. The best experimental data available is that taken at 1,046 Mc by the NBS group [72]. The statistical problem is very similar to that illustrated in figure 26, if one reinterprets R_1 as the instantaneous signal amplitude induced in the first receiver and R_2 that induced in the second, at the same instant. The joint probability density for R_1 and R_2 is given by eq (4.6.7), if the basic correlation coefficient $\mu(D)$ is now the space correlation between the orthogonal signal components at the two receivers. Staras [73] computed the probability that the greater of the two (diversity) signals is above a specified level, as a function of the correlation coefficient $\mu\,(D)$. These statistical predictions have been checked against experimental data [74] and seem to fit quite well.

The propagation problem is to estimate the space coefficient $\mu\,(D)$. Gordon [63] and Rice [68] presented qualitative analysis which showed that the correlation distance normal to the propagation path

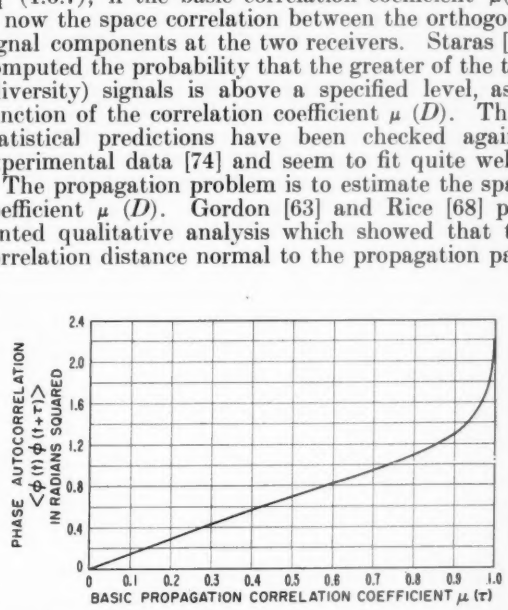


FIGURE 26. *Autocorrelation of scatter signal phase versus basic propagation correlation coefficient* $\mu\,(\tau)$.

[10] R. J. Wagner, to be published.

(horizontal and vertical) ought to be of the order $\lambda a/d$, where d is the transmission distance and a the radius of the earth. A correlation distance $\lambda a^2/d^2$ along the path was also estimated. The RCA experiments [74] at 390 Mc for path lengths of 73 and 153 miles, however, failed to show the predicted decrease of horizontal and vertical correlation lengths with path distance. On the other hand, the simple wavelength dependence of these results is borne out by multiple frequency experiments [53, 75].

Analytical propagation calculations of all three correlation functions[11] as functions of receiver separation distance were derived by Staras [64], using the exponential space correlation model. His results have the same form as Gordon's, but are numerically different by substantial factors. Staras also included the effects of anisotropy in the refractive irregularities and found that this affected the predicted horizontal (transverse) correlation distance,

$$S=\frac{1.4\pi}{r}\left(\frac{a}{d}\right)\lambda, \tag{4.8.1}$$

where r is the vertical-to-horizontal scale length ratio. By fitting data taken on 100 and 1000 Mc [53, 72] he suggested that this ratio should be about 0.25. Such comparisons are important in that they bear directly on the question of anisotropy, but more good data is needed.

The realization of plane-earth height-gain advantages is limited by space coherence between the actual receiver and its image. The usual height-gain curve is realized if the antenna is less than one-half the vertical correlation distance ($\lambda a/2d$). Only slight improvement is gained above this height by the gradual decrease of the scattering angle.

4.9. Frequency Correlation and Medium Bandwidth

The signal-frequency bandwidth which the scatter medium can support without serious distortion is intimately related to the multipath delays which are experienced. These depend upon the size of the effective scatter volume, which in turn is determined by the antenna bandwidths and/or the scattering pattern of the blobs. When the antenna patterns (β) are broad, the multipath is controlled by the scattering blobs and the approximate result of Gordon [63] predicts that

$$\Delta f=\frac{5ca^2}{d^3}, \qquad \beta>>\frac{d}{a} \tag{4.9.1}$$

or numerically

$$\Delta f_{\text{Mc}}=30/d^3,$$

where d is in hundreds of miles. This result shows that the bandwidth decreases rapidly with distance. Rice's estimate [68] of Δf varies inversely with the path length d, but is based on an oversimplified physical model.

A real improvement in bandwidth capabilities is obtained with very narrow beams. Although some antenna gain degradation is experienced with narrow beams, the angular distribution of multipath signals which are admitted to the receiver is reduced and the bandwidth thereby increased. Booker and de Bettencourt [76] examined this problem for beams narrower than the average scattering angle $\bar{\theta}=d/a$ and found

$$\Delta f=\frac{4ca}{\beta d^2}, \qquad \beta<<\frac{d}{a} \tag{4.9.2}$$

This result indicates steady improvement with antenna size and decreases less rapidly with distance than (4.9.1). When aircraft fly through the path, the solid bistatic reflections from them supply a multipath signal of unusual magnitude and the above estimates for broad-antenna beams are meaningless.

Figure 27 indicates typical successive plots of amplitude versus frequency, such as might be obtained by sweeping a narrow-band transmitter about the carrier f_0. The joint amplitude distribution (4.6.7) can be used to describe this situation if R_1 is interpreted as the signal strength at f_0 and R_2 that at $f_0+\Delta f$. The basic correlation function μ is now a function of the frequency separation Δf and other propagation variables. If the correlation is high, it means that the signal components at either end of the frequency band being used are behaving essentially in unison, and little or no distortion is expected. Staras [64] evaluated substantially this function directly for omnidirectional antennas with

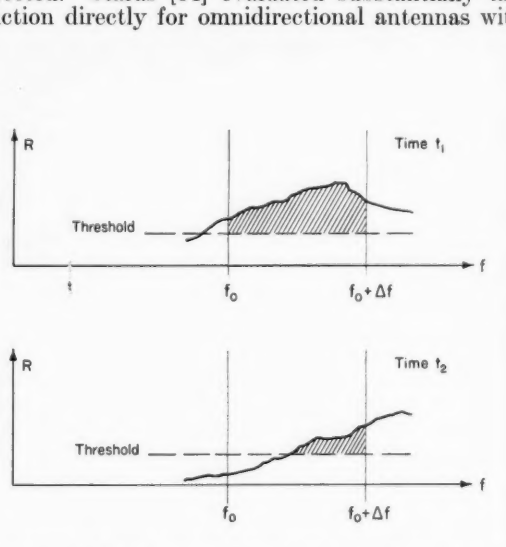


FIGURE 27. *Typical successive amplitude records plotted versus carrier frequency over a band Δf.*

The receiver threshold level indicates the role of absolute power level in correlation comparisons.

[11] Along, horizontally normal (transverse), and vertically normal to the propagation path.

the exponential space correlation model, and finds that the frequency correlation drops to 0.5 when

$$\Delta f_{mcs}=\frac{4}{d^3} \tag{4.9.3}$$

where d is in hundreds of miles. Reference to figure 24 indicates that (4.9.3) refers also to the envelope (R) correlation, so a factor of eight difference exists between (4.9.1) and (4.9.3). This problem evidently requires further clarification and careful comparison with experimental data. It is possible to use the spectral separation of turbulence and propagation effects indicated by (4.6.4) to calculate the frequency correlation coefficient $\mu(\Delta f)$. One need only evaluate the results of the frequency-dependent geometric integrals at displaced frequencies and set $\tau=0$ in (4.6.4).

The published experimental data has been taken with narrow-antenna beams, and is therefore not directly useful in resolving the theoretical conflict. The recent Lincoln Laboratory data [77] measures the frequency correlation in the manner of figure 27, by using a rapid frequency-change system. Their results indicate that signals separated by 2 Mc at 2,300 Mc over a 188-mile path give an amplitude correlation of 0.4. This is consistent with the television experiments of Tidd [78] on a similar path at 5,000 Mc.

It is important to note that the receiver threshold level influences the apparent bandwidth of a system. Figure 27 indicates how amplitude fading at time t_2 can reduce the frequency band over which the signal is correlated with the proceeding signal. This has lead to the concept of "instantaneous bandwidth," which finds application in designing efficient modulation systems to work with the scatter propagation mode.

4.10. Gain Loss or Antenna-to-Medium Coupling

When very narrow beams are used on scatter circuits, it is found that free-space antenna gains are not realized. This is known as gain loss or antenna-to-medium coupling loss and arises from the fact that signals arrive at the receiver from an extended scattering volume V. Narrowing the antenna beam eventually reduces this common volume faster than the aerial gain is increased (see fig. 16b) so that a relative gain loss is incurred.

Booker and de Bettencourt [76] analyzed this problem for similar conical antenna patterns at the transmitting and receiving ends of the circuit. They argued that no gain loss should be incurred so long as the beamwidth β is greater than $\frac{2}{3}d/a$. When the beams are sharper than this average scattering angle, they suggest that a relative power loss factor

$$L=0.43\left(\frac{d}{\beta a}\right)^3, \tag{4.10.1}$$

will be incurred. They point out that this result is probably too large because the height dependence of $\langle\Delta\epsilon^2\rangle$ was neglected and sharp beams were employed. The calculations by Nortion, Rice, and Vogler [58] used an exponential correlation function and an inverse square height dependence of $\langle\Delta\epsilon^2\rangle$. Staras [79] treated anisotropic irregularities and dissimilar antenna patterns. His results are between 5 and 8 db less severe than these predictions by Booker and de Bettencourt, and seem to give rather good agreement with experimental data.

Hartman and Wilkerson [80] have reinvestigated the problem, using an exponential decrease of isotropic turbulent fluctuations with height. Their calculations produce gain loss estimates which are several decibels less than those predicted by Staras, and in good agreement with considerable experimental data.

4.11. Beam Swinging

It is found experimentally that the scatter signal amplitude decreases when narrow beamwidth antennas are jointly swung off the great circle path, or simultaneously elevated as shown in Figure 28. For azimuth swinging, the effect is simply to increase the average scattering angle θ between the ray paths (beams) [81]. This reduces the scattered power because of the large exponent of $\sin\theta/2$ in the various scattering cross section expressions (see sec. 4.2).

Booker and de Bettencourt [76] analyzed the effect of swinging the transmitting and receiving aerials simultaneously in azimuth and elevation. They used a height-independent exponential correlation to make estimates of this power reduction, which are apparently in good agreement with available data.

The dependence of power loss on elevation swinging must also recognize the observed decrease of turbulent intensity with height. Since $\langle\Delta\epsilon^2\rangle$ probably decreases exponentially with height, the concomitant scattering height change Δh can effect the power levels significantly if it is comparable with the atmospheric scale height $H=22{,}000$ ft. The problem evidently needs further analysis and comparison of theory and experiment.

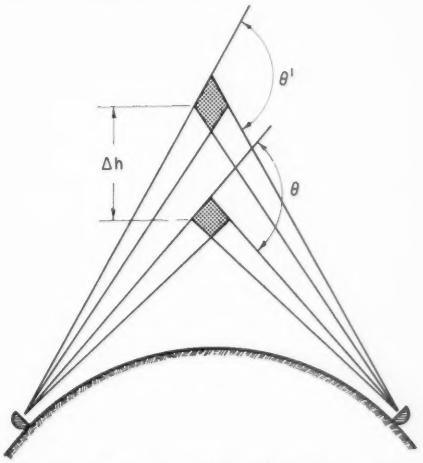


FIGURE 28. *Geometry for vertical beam swinging experiment showing increased scatter angle and higher common volume.*

The preparation of this paper is due in large part to the hospitality extended by the Boulder Laboratories of the NBS and to invaluable discussions with my colleagues there. W. Hartman and E. Barrows of that laboratory were particularly helpful in bringing the paper to final form and checking the analytical results.

5. References

[1] C. M. Crain, Survey of airborne microwave refractometer measurements, Proc. I.R.E. **43,** 1405 (1955).

[2] G. Birnbaum and H. E. Bussey, Amplitude, scale and spectrum of refractive index inhomogeneities in the first 125 meters of the atmosphere, Proc. I.R.E. **43,** 1412 (1955).

[3] M. C. Thompson and H. B. Janes, Measurements of phase stability over a low-level tropospheric path, J. Research NBS **63D,** 45 (1959).

[4] Albert D. Wheelon, Relation of radio measurement to the spectrum of tropospheric dielectric fluctuations J. Appl. Phy. **28,** 684 (1957).

[5] D. Mintzer, Wave propagation in a randomly inhomogeneous medium, J. Acoustical Soc. Am. I: **25,** 922 (1953), II:**25,** 1107 (1953), III:**25,** 186 (1954).

[6] H. G. Booker and W. E. Gordon, The role of stratospheric scattering in radio communication, Proc. I.R.E. **45,** 1223 (1957).

[7] F. Villars and V. F. Weisskopf, The scattering of electromagnetic waves by turbulent atmospheric fluctuations, Phys. Rev. **94,** 232 (1954).

[8] F. Villars and W. F. Weisskopf, On the scattering of radio waves by turbulent fluctuations of the atmosphere, Proc. I.R.E. **43,** 1232 (1955).

[9] A. M. Obukhov, Structure of the temperature field in turbulent flow, Izvest. Akad. Nauk S.S.S.R. Ser. Geofiz. **13,** 58 (1949).

[10] R. A. Silverman, Turbulent mixing theory applied to radio scattering, J. Appl. Phys. **27,** 699 (1956).

[11] R. Bolgiano, The role of turbulent mixing in scatter propagation, IRE Trans. PGAP **6,** 159 (1958).

[12] Albert D. Wheelon, Spectrum of turbulent fluctuations produced by convective mixing of gradients, Phys. Rev. **105,** 1706 (1957). (See also J. Geophys. Research **62,** 93 (1957), Radio frequency and scattering angle dependence of ionospheric scatter propagation at VHF, et seq.)

[13] B. R. Bean and F. M. Meaney, Applications of the monthly median refractivity gradient in tropospheric propagation, Proc. I.R.E. **43,** 1419 (1955).

[14] R. A. Silverman, Fading of radio waves scattered by dielectric turbulence, J. Appl. Phys. **28,** 506 (1957).

[15] W. Heisenberg, Zur statistischen theorie der turbulenz, Z. physik. **124,** 628 (1948).

[16] C. F. Von Weizsacker, Das spektrum der turbulenz bei grossen Reynolds'schen zahlen, Z. physik. **124,** 614 (1948).

[17] G. Munch and A. D. Wheelon, Space-time correlations in stationary isotropic turbulence, Phys. Fluids **I,** 462 (1958).

[18]. S. O. Rice, Mathematical analysis of random noise, BSTJ, **24,** 46, 100 et seq. (1945).

[19] K. A. Norton, L. E. Vogler, W. V. Mansfield, and P. J. Short, The probability distribution of the amplitude of a constant vector plus a Rayleigh-distributed vector, Proc. I.R.E. **43,** 1354 (1955).

[20] H. Bremmer, On the theory of fading properties of a fluctuating signal imposed on a constant signal, NBS Circ. 599 (May 1959).

[21] K. A.Norton, E. L. Schultz, and H. Yarbrough, The probability distribution of the phase of the resultant vector sum of a constant vector plus a Rayleigh-distributed vector, J. Appl. Phys. **23,** 137 (1952).

[22] J. R. Johler, and L. C. Walters, The mean absolute value and standard deviation of the phase of a constant vector plus a Rayleigh-distributed vector, J. Research NBS **62,** 183 (1959) RP2950.

[23] Albert D. Wheelon, Near-field corrections to line-of-sight propagation, Proc. I.R.E. **43,** 1459 (1955).

[24] M. Balser, Some observations on scattering by turbulent inhomogeneities, IRE Trans. PGAP **AP-5,** 383 (1957).

[25] S. Stein, Some observations on scattering by turbulent irregularities, IRE Trans. PGAP **AP-6,** 299 (1958).

[26] T. H. Ellison, The propagation of sound waves through a medium with very small random variations in refractive index, J. Atmospheric and Terrest. Phys. **2,** 14 (1951).

[27] S. M. Rytov, Light diffraction by ultrashort waves, Izvest. Akad. Nauk S.S.S.R. Ser. Fiz. No. 2 (1937).

[28] J. Feinstein, Some stochastic problems in wave propagation—Part II, IRE Trans. **AP-2,** 63, (1954).

[29] R. B. Muchmore, and A. D. Wheelon, Line-of-sight propagation phenomenon, I Proc. I.R.E. **43,** 1437 (1955).

[30] J. W. Herbstreit, et al., Radio studies of atmospheric turbulence, vol. I and II, May 31, 1956 (to be published).

[31] E. Levin, R. B. Muchmore and A. D. Wheelon. Aperature-to-medium coupling on line-of-sight paths: Fresnel scattering, to be published in IRE Trans. PGAP (1959).

[32] K. A. Norton, P. L. Rice, H. B. Janes and A. P. Barsis, The rate of fading in propagation through a turbulent atmosphere, Proc. I.R.E. **43,** 1341 (1955).

[33] J. W. Herbstreit, and M. C. Thompson. Measurements of the phase of radio waves received over transmission paths with electrical lengths varying as a result of atmospheric turbulence, Proc. I.R.E. **43,** 1391 (1955).

[34] A. P. Dean, and B. M. Fannin, Phase difference variations in 9350 mc radio signals arriving at spaced antennas, Proc. I.R.E. **43,** 1402 (1955).

[35] E. Barrows, Effect of tropospheric fluctuations on radio star signals, to be submitted for publication to J. Research NBS, Section D (Radio Propagation).

[36] V. A. Krasil'nikov, The effect of variations of the coefficient of refraction in the atmosphere upon the propagation of ultrashort radio waves, Izvest. Akad. Nauk. S.S.S.R. Sev. Geograf i Geofiz **13,** 33 (1949).

[37] P. G. Bergman, Propagation of radiation in a medium with random inhomogeneities, Phys. Rev. **70,** 486 (1946).

[38] E. C. S. Megaw, Interpretation of stellar scintillation, Quart. J. Royal Meteorol. Soc. **80,** 248 (1954).

[39] A. M. Obukhov, On the effect of weak atmospheric inhomogeneities on sound and light propagation, Izvest. Akad. Nauk. S.S.S.R. Ser. Geofiz. No. 2, 155 (1953).

[40] L. A. Chernov, Correlation of the amplitude and phase fluctuations in wave propagation in media with random inhomogeneities, Akusticheskii Zhurnal **1,** 89 (1955) and A. N. Doklady, USSR **98,** 953 (1954).

[41] V. I. Tatarskii, On the theory of the propagation of sound waves in a turbulent stream, Zhur. Eksptl'i i Theoret Fiz. **25,** 74 (1953).

[42] B. M. Fannin, Line-of-sight wave propagation in a randomly inhomogeneous medium, IRE Trans. PGAP, **AP-4,** 661 (1956).

[43] R. P. Feynman, Space-time approach to quantum electrodynamics, Phys. Rev. **76,** 769 (1949).

[44] R. B. Muchmore, and A. D. Wheelon, Line-of-sight propagation phenomenon II, Proc. I.R.E. **43,** 1437 (1955).

[45] C. L. Pekeris, Note on the scattering of radiation in an inhomogenous medium, Phys. Rev. **71,** 268 (1947).

[46] H. G. Booker, and W. E. Gordon, A theory of radio scattering in the troposphere, Proc. I.R.E. **38,** 401 (1950).

[47] See papers in the Scatter Prop. Issue, Proc. I.R.E. **43** (Oct. 1955).

[48] K. Bullington, Characteristics of beyond-the-horizon radio transmission, Proc. I.R.E. **43,** 1175 (1955).

[49] H. Staras, and A. D. Wheelon, Theoretical research on tropospheric scatter propagation in the United States 1954–1957, IRE Trans. PGAP **AP–7,** 80 (1959).

[50] R. A. Silverman, Some remarks on scattering from eddies, Proc. I.R.E. **43,** 1253 (1955).
[51] M. Balser, and R. A. Silverman, Statistics of electromagnetic radiation scattered by a turbulent medium, Phys. Rev. **96,** 960 (1954).
[52] H. B. Janes, An analysis of within-the-hour fading in 100-to-1000-Mc transmissions, J. Research NBS **54,** 231 (1955) RP 2585.
[53] A. P. Barsis, J. W. Herbstreit and K. O. Hornberg, Cheyenne Mountain tropospheric propagation experiments, NBS Circ. 554 (1955).
[54] G. McCrossen, Analysis of the long term distribution of instantaneous received power (to be published).
[55] L. D. Arons, An analysis of radio-wave scattering in the diffraction region, Cornell Univ. E. E. Report 312 (30 Oct. 1956, Ithaca, N.Y.).
[56] H. Bremmer, Terrestial Radio Waves (Elsevier Press, Inc., Houston, Texas, 1949).
[57] H. Feshbach, Unpublished Lincoln Laboratory Memoranda.
[58] K. A. Norton, P. L. Rice, and L. E. Vogler. The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. I.R.E. **43,** 1488 (1955).
[59] J. H. Chrisholm, P. A. Portmann, J. T. deBettencourt, and J. F. Roche, Investigations of angular scattering and multipath properties of tropospheric propagation of short radio waves beyond the horizon, Proc. I.R.E. **43,** 1317 (1955).
[60] B. R. Bean, and R. Abbott. Oxygen and H_2O vapor absorption of radio waves in the atmosphere, Geofis pura e appl, **37,** 127 (1957).
[61] R. L. Abbott, unpublished work.
[62] P. L. Rice, A. G. Longley, and K. A. Norton, Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss, NBS Technical Note 15 (July 1959).
[63] W. E. Gordon, Radio Scattering in the Troposphere, Proc. I.R.E. **43,** 25 (1955).
[64] H. Staras, Forward scattering of radio waves by anisotropic turbulence, Proc. I.R.E. **43,** 1374 (1955).
[65] T. F. Rogers, VHF field strength far beyond the radio horizon, Proc. I.R.E. **43,** 623 (1955).
[66] J. W. Herbstreit, K. A. Norton, P. L. Rice, and G. E. Schafer, Antennas and propagation, IRE Convention Record, Part 2, pp. 85–93 (March 1953).
[67] A. T. Waterman, Some generalized scattering relationships in transhorizon propagation, Proc. I.R.E. **46,** 1842 (1958).
[68] S. O. Rice, Statistical fluctuations of radio field strength far beyond the horizon, Proc. IRE **41,** 274 (1953).
[69] J. B. McGuire, and A. D. Wheelon, Calculation of the fading rate for tropospheric scatter propagation, Presented at spring URSI Meeting, Washington, D.C. (1958).
[70] V. I. Bunimovich, Fluctuating processes as oscillations with random amplitudes and phases, Zhur. Tekh. Fiz. **19,** 1231 (1949).
[71] C. L. Mack, Diversity Reception in UHF Long Range Communications, Proc. I.R.E. **43,** 1281 (1955).
[72] A. F. Barghausen, M. T. Decker, and L. J. Maloney, Measurements of correlation, height gain, and path antenna gain at 1046 Mc on spaced antennas far beyond the radio horizon, 1955 IRE Convention Record, Part I, Vol. 3.
[73] H. Staras, Diversity reception with correlated signals, J. Appl. Phy. **27,** 93 (1956), see also Proc. I.R.E. **44,** 1057 (1956).
[74] R.C.A. Report, Study and investigation of tropospheric scattering, Part A, **91** (1956), per U.S. Army Signal Corps Eng. Labs, Contract DA–36–039–SC–64555.
[75] H. B. Janes and P. I. Wells, Some tropospheric scatter propagation measurements near the horizon, Proc. I.R.E., **43,** 1336 (1955).
[76] H. G. Booker and J. T. de Bettencourt, Theory of radio transmission by tropospheric scattering using very narrow beams, Proc. IRE, **43,** 281 (1955).
[77] J. H. Chisholm, L. P. Rainville, J. F. Roche, and H. G. Root, Measurements of the bandwidth of radio waves propagated by the troposphere beyond the horizon, IRE Trans. PGAP, **AP–6,** 377 (1958).
[78] W. H. Tidd, Demonstration of bandwidth capabilities of beyond-the-horizon tropospheric radio propagation, Proc. IRE., **43,** 1297 (1955).
[79] H. Staras, Antenna-to-medium coupling loss, IRE Trans. PGAP, **AP–5,** 288 (1957).
[80] W. Hartman and R. Wilkerson, Path antenna gain in an exponential atmosphere, to be published in J. Research NBS, 63D (Nov. 1959).
[81] D. K. Bailey, R. Bateman, and R. C. Kirby, Radio transmission at VHF by scattering and other processes in the lower ionosphere, Proc. IRE, **43,** 1181 (1955).

BOULDER, COLO. (Paper 63D2–19.)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Study at 1,046 Megacycles per Second of the Reflection Coefficient of Irregular Terrain at Grazing Angles

Raymond E. McGavin and Leo J. Maloney

(February 12, 1959)

An experimental determination of the reflection coefficient over rough terrain is reported. The reflected signal received over rough terrain is considered to be made up oft wo components, one that is a specular component and the other a Rayleigh-distributed component. Where one terminal is low, the Rayleigh component is considered to be small with respect to the specular component but increases in relative magnitude as the height of the lower terminal increases. A terminal height is reached where the specular component is no longer significant, and the reflected energy is essentially Rayleigh-distributed. A terminal height is quickly reached above which the mean value of the reflected energy is relatively constant, of a low value, and independent of the grazing angle.

1. Introduction

The interest in rough-terrain reflection has led to the publication of a large number of papers on this subject in recent years. The complexity of the problem is evidenced by the many different approaches to reasonable and usable solutions. The most obvious approach to the problem would be to consider rough-terrain reflection as a modification of smooth-terrain reflection (i.e., specular reflection) and attempt to express the rough-terrain reflection coefficient in terms of the smooth earth or Fresnel reflection coefficient. However, the problem is not at all that simple.

The field measured at a point in space above the earth and within the radio line-of-sight of the transmitter is the vector sum of a number of signal components arriving at the receiver over a number of different paths. In general, there is a single wave that arrives directly from the transmitter, while others arrive after being reflected at least once from the intervening ground. Ideally, where the ground is a smooth surface, modified image theory adequately describes the received field, which is the vector sum of a direct wave and a single ground-reflected wave [1].[1] Where the ground contains known irregularities, few in number, and where the reflecting area is small, a reasonable estimate of the reflected energy can be determined on a Fresnel diffraction basis [2]. However, if the irregularities are numerous and haphazard and if the reflecting area is extensive, it is very difficult to determine Fresnel zones.

Some methods of treatment have been proposed [3 through 10] for a rough surface where an exhaustive knowledge of the surface is available, or where the surface is such that it can be approximated by a simple function or where the irregularities are small with respect to the wavelength. In the absence of such idealized conditions, an analytical determination of the coefficient of reflection is not possible, and it has been common practice to assume specular reflection even over rough terrain when the reflection occurs near grazing relative to an average smooth ground.

Any successful theory would have to explain the inherent relationships between the pertinent parameters. The authors of this paper do not attempt to propound a theory to explain the physics of reflection from extensive randomly rough terrain having large irregularity, but do attempt to point out some relationships between these observed parameters. The approach must be statistical since it can be shown that, for such surfaces only this method leads to practical application [7]. Attention will be concentrated on grazing angles which are of greatest interest. The angle between the incident wave and the tangent to an approximating smooth surface at the point of reflection will be called the grazing angle if it is less than 5°.

2. Experimental Procedure

During August of 1954, a series of aircraft flights were made over transmission paths in eastern Colorado and western Kansas. Three paths were investigated: one originating at the Cheyenne Mountain transmitting site near Colorado Springs, Colo. passing through Haswell, Colo. and continuing beyond; another from Pikes Peak through Haswell; and a third from Fort Carson through Haswell. The mean ground elevation of these paths varies from 6,200 ft relative to mean sea level near Cheyenne Mountain to 4,200 ft in western Kansas. The object of the flights was to investigate the distribution of received field strengths along these paths at an operating frequency of 1,046 Mc using horizontal polarization. A complete description of the transmission facilities can be found elsewhere [11]. Table 1 shows the pertinent data concerning the trans-

[1] Figures in brackets indicate the literature references at the end of this paper.

mitting and receiving system. All flights with the exception of the 7,000-ft and the 10,000-ft flights on the Cheyenne Mountain path were flown in a west-to-east direction. For the flights in the opposite direction, any difference in the antenna pattern and gain due to the required change in mounting was ignored, since absolute values of the field or even values of the field relative to each other for any two flights were not required. Diurnal effects in the characteristics of transmission have also been neglected since such effects would be expected to influence the values of reflection coefficient obtained only slightly. Likewise, standard air refraction is assumed; although the meteorological data collected at the time indicate some departure from a standard atmosphere, this also would not be expected to affect our conclusions appreciably.

Visual checkpoints were used to maintain the aircraft on the path and to locate the aircraft over the terrain. For purposes of calculation, it was assumed that the aircraft maintained a constant speed between checkpoints. This assumption seems warranted as the average speed for successive sectors did not vary by more than 5 percent.

For the most part, the entire region over which these measurements were made is very dry and is considered to be semiarid. A major portion of the ground is covered with Buffalo grass and sagebrush; a small part is under cultivation. The entire region is essentially without trees. Although there are isolated ranches and farms, there are no areas of concentrated population along the paths.

3. Analysis

3.1. Interpretation of the Reflecting Surfaces of the Flight Paths

Inasmuch as Rayleigh's criterion of roughness has been generally accepted as an estimate of the effective boundary between rough and smooth terrain, this criterion has been used in the following analysis. Rayleigh's criterion of roughness,

$$\Delta\phi=\frac{4\pi}{\lambda}\Delta h \sin\psi,$$

is an expression relating the root mean square deviations of the surface from a smooth surface (Δh), the wavelength of the signal (λ) and the grazing angle (ψ). This phase difference ($\Delta\phi$) of rays reflected from different levels of a surface determines the effective surface roughness. Critical values for this expression ranging from $\pi/2$ to 0.1 have been used by various authors [12, 13, 14]. Beckman [7] uses this criterion to determine the existence of a horizontal reflecting element in a profile. He also uses a modification of the same expression to determine the classification of the surface reflection, from one that is characterized by the Dirac distribution as one extreme, to one that is characterized by the Rayleigh distribution as the other extreme. In any case, irrespective of the particular limiting value chosen, if only the area within the smooth-earth first Fresnel zone is investigated, all paths considered here must be classified as being rough at least for grazing angles greater than 0.4°, and for frequencies in excess of 1,000 Mc. It should be noted, however, that there is not exact correspondence between Beckman's consideration and that stated here. Beckman would consider a detailed analysis of the entire profile and the selection of horizontal elements. This method considers the average deviations from what is considered to be the principal reflecting area. A detailed surface analysis is not practical for areas as large as those involved in this study.

Another problem in the determination of the reflection coefficient is the consideration of divergence. When a wave is reflected from a curved smooth surface, the power density is decreased due to the spreading of the reflected wave. As the angle of the incident wave approaches grazing over a smooth curved surface, divergence causes the power density of the reflected wave to decrease to a very-low value. However, if the terrain is sufficiently rough, it is difficult to determine the effect of divergence, since

TABLE 1. *Transmitting and receiving system data*

Site location	Transmitter						Receiver			
	Elevation		Power output	Antenna			Flight levels		Antenna	
	Above mean sea level	Above average terrain		Type	Gain [a]	Vertical beam width	Above mean sea level	Average height over terrain[b]	Type	Gain [a,c]
	ft	ft					ft	ft		
Pikes Peak	14,110	7,491	18 w	Horn-fed parabolic section	24 db	6°	10,000	4,650	Yagi	6 db
							7,000	1,650		
Cheyenne Mountain	8,760	2,836	4 kw	Slotted waveguide fed horn.	26 db	6°	15,000	9,600	Yagi	6 db
							10,000	4,600		
							7,000	1,600		
Fort Carson	6,250	35	18 w	Dipole-fed parabola	22 db	[c] 8°	15,000	9,150	Yagi	6 db
							10,000	4,150		

[a] Gains are relative to an isotropic radiator.
[b] Calculated near the center of the path.
[c] Estimated.

Note: Operating Frequency: 1,046 Mc
Polarization: Horizontal
Receiver: APR-5 aircraft
Aircraft: B-17

it would be reasonable to expect anomalies in the region of the radio horizon which would be more effective in determining the amount of reflection than would the general curvature of the surface at that point. Hence, the reflection coefficient spoken of here could be called the "effective reflection coefficient" as it has not been corrected for possible divergence caused by reflection from the curved surface of the earth.

In order to relate the results to a reflecting region on the ground, curves were fitted to the terrain and for the purposes of calculation of the grazing angles, the optical reflection point of the surface was assumed to lie on the curve rather than on the actual terrain [12]. In the case of the Pikes Peak and the Cheyenne Mountain paths, a single second order curve was drawn through the entire profile, and it was assumed that the curve was an approximation of the terrain, at least in a macroscopic sense. Since the terrain over the Fort Carson path did not permit a simple geometrical approximation, the results for the Fort Carson flights were related to the angle of elevation of the receiver above the transmitter radio horizon, rather than to the grazing angle.

3.2. Characteristic Effects of Specular and Random Reflections

When reflection from a smooth surface is considered, principles of optical reflection can be applied and the reflection can be viewed as though it occurred at a point. Then the reflection coefficient is the ratio of the electric field reflected to the electric field incident, considered along paths where the angle of reflection equals the angle of incidence. This reflection coefficient has both an amplitude and a phase, both of which are functions of the ground constants and the angle that the incidence ray makes with the tangent to the surface at the optical point of reflection. If it is assumed that the direct and ground-reflected waves arriving at the receiver left the transmitter with the same power density and traveled over paths approximately equal in length, the ratio of the amplitude of the reflected wave to the amplitude of the direct wave is equal to the amplitude of the reflection coefficient. Where the terrain is reasonably smooth, a well-defined "lobe structure" of field maximums and minimums is formed in space due to the interference between the direct and the ground-reflected waves. The accepted practice in evaluating the amplitude of a specular reflection coefficient over smooth terrain is to consider successive maximums and minimums of the field strengths of the received signal as a function of distance and to evaluate the reflection coefficient from the following expression:

$$\rho=\frac{d_2}{d_1}\frac{\left(\frac{E_M}{E_m}\right)-1}{\left(\frac{E_M}{E_m}\right)+1}, \qquad (1)$$

where:

d_1 = distance traveled by the direct wave,
d_2 = distance traveled by the reflected wave,
E_M = the value of a voltage maximum, and
E_m = the value of a preceding or following voltage minimum.

The well-defined maximums and minimums lend themselves well to the above method of analysis, as the reflected wave can be considered to be a single ray. However, where the reflecting surface is very rough, such a model lobe structure does not exist. A moving receiver would experience fading at a more or less haphazard rate, and the validity of the above treatment is questionable. Since it is desirable to define the reflection coefficient over rough terrain in a manner similar to that over smooth terrain, the determination of the reflection coefficient should be in terms of grazing angles similar to those over the average smooth earth. If a least-square curve is substituted for the terrain and optical reflection points are assigned by geometry, the reflection coefficient can be defined as the electric field reflected to that incident along paths where the angle of incidence equals the angle of reflection over the average smooth earth. It should be remembered, however, that the reflected energy is not arriving via a single path anymore than it is in the case of specular or smooth-earth reflection as can be shown from a Fresnel diffraction treatment [2] of smooth plane-surface reflection. However, with rough terrain the problem is somewhat different since the reflected field becomes incoherent. Image theory is not adequate here, and the reflected energy, rather than being considered to consist of an essentially single reflected ray, must be considered more completely as consisting of the vector sum of many waves reflected from an extensive surface. The energy is diffused over a large region of space upon reflection and the reflected energy arriving at the receiver must be considered to be arriving from many different points on the reflecting surface at many different phase angles. The vector sum of these, if it is assumed that no one of the vectors is significantly large with respect to any of the others, would be a vector whose relative phase would be random and whose amplitude, E_s, would be distributed in accordance with the Rayleigh distribution [15]:

$$\Pr\,(E_s>Z)=\exp\,(-Z^2/k^2) \qquad (2)$$

where $k^2=E_1^2+E_2^2+\ \ldots\ +E_n^2=\Sigma_{i=1}^{n}E_i^2=$constant.

Then over rough terrain the signal arriving at the receiver can be interpreted as being the sum of a constant vector (the direct wave) and a Rayleigh distributed vector (the reflected wave). Perhaps this may seem to be a simplification of the problem in that it assumes that the surface reflection is characterized by a single distribution, namely the Rayleigh distribution. As was previously pointed out, at least one author [7] considers the Rayleigh distribu-

tion to be the limiting case in so far as scattering in the principal direction (the direction in which specular reflection would take place if the surface were smooth) is concerned; i.e., it is required that the surface irregularities be quite large with respect to the wavelength. The two approaches are not wholly inconsistent, as it was pointed out that the departures of this terrain from a smooth surface are considerable even at small grazing angles.

3.3. Separation of a Constant Vector and a Rayleigh-Distributed Vector

The sum of a direct wave and a wave reflected from a randomly rough surface will be considered to be composed of a constant component and a Rayleigh-distributed component, and can be separated into these two components by a consideration of the slope of the cumulative distribution of the received signal strength over discrete intervals of distance along the path [15, 16]. The probability distribution of the amplitude (r') of the sum of a constant vector plus a Rayleigh distributed vector was considered by Norton, Vogler, Mansfield, and Short [16], based on the expansion of a probability distribution derived by Rice [17] in the analysis of random noise:

$$\Pr\,(r'>r)=\frac{2}{k^2}\int_r^\infty r\exp\,[-(1+r^2)/k^2]I_0(2r/k^2)dr, \quad (3)$$

where r'=instantaneous resultant amplitude, and where

$$I_0(x)=\frac{1}{\pi}\int_0^\pi \exp\,(x\cos\phi)d\phi.$$

In figure 1, a relationship is given between the fading range in decibels R(0.1)—R(0.9) and the level K in decibels of the energy of the reflected wave relative to that of the direct wave. For example, when the ratio of the 10 percent to 90 percent values of the distribution of the received field over a discrete distance interval was 4.9 db, the average energy in the reflected wave would be expected to be 10 db below that of the direct wave and this corresponds to an effective rms amplitude of the reflection coefficient of 0.316. From such consideration, it can be seen that the reflection coefficient as defined above is expressed by a distribution and can have many different values. Any expression of a particular value must be made in terms of its statistical probability of occurrence. In this example 0.316, the rms amplitude of the reflection coefficient, is that value which would be exceeded by 36.8 percent of all the possible values that the reflection coefficient may have; whereas the median value of the same reflection coefficient would be 0.263 which would be the value exceeded by 50 percent of all the values to be expected.

This theory, as can be seen from eq (2), assumes that the average energy of the reflected wave does not change appreciably from point to point within the distance intervals chosen. The distance intervals

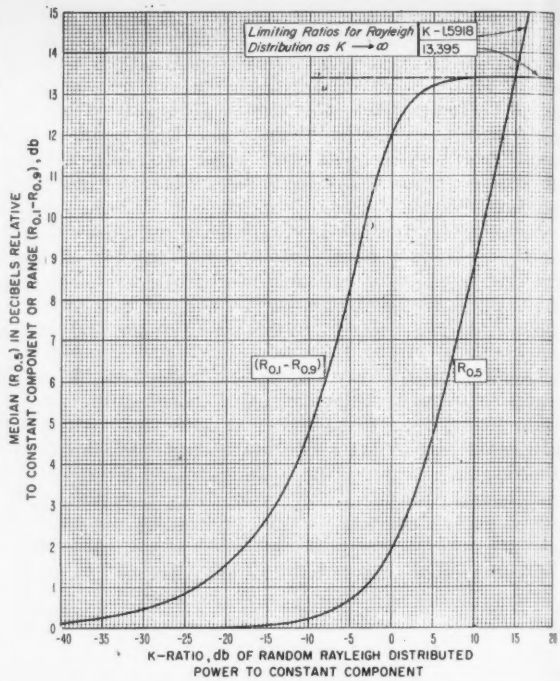

Limiting Ratios for Rayleigh Distribution as K → ∞
K - L5918
13.395
MEDIAN (R0.5) IN DECIBELS RELATIVE TO CONSTANT COMPONENT OR RANGE (R0.1-R0.9), db
(R0.1 - R0.9)
R0.5
K-RATIO, db OF RANDOM RAYLEIGH DISTRIBUTED POWER TO CONSTANT COMPONENT

FIGURE 1. *Median* ($R_{0.5}$) *and range* ($R_{0.1}$—$R_{0.9}$) *from the cumulative distribution of the resultant amplitude of a constant vector plus a Rayleigh-distributed vector.*

Power in random component is K db relative to the range ($R_{0.1}$—$R_{0.9}$) of the cumulative distribution.

were chosen somewhat arbitrarily to compromise the conflicting needs for sufficiently homogeneous data and sufficiently small sampling error. Intervals of one mile were chosen except close to the radio horizon where three, five and even ten mile intervals were used. For each interval, cumulative distributions of signal level versus distance were plotted on log-Rayleigh coordinates and the 10 percent and 90 percent values were found. The value K was found from figure 1 for each value of R(0.1)—R(0.9), and since K=20 log k, each interval produced an estimate of the parameter k of a Rayleigh distribution; that is, an estimate of the rms amplitude of the reflection coefficient. In the region of the radio horizon, as the grazing angle approaches zero, the spatial fading rate decreases and a quasi-lobe structure is formed. In this region it is easy to compare the two methods. For each interval, two values were determined; the rms value of several determinations of the reflection coefficient as determined considering successive maximums and minimums, and the rms value of a Rayleigh distribution for the interval assuming ground reflection to be random. The two different methods were then compared by use of the Wilcoxon signed rank test [18, p. 182]. Only two of the seven flights indicated statistically significant difference in the two methods. This comparison is mentioned not to question the validity of the approach presented here, but to point out its

consistency with the conventional method in the interval near the horizon. It must be remembered that the conventional method assumes that the reflection coefficient does not change, at least in the interval between successive maximums and minimums. This requirement is more restrictive than the assumption that the average reflected energy does not change from period to period within the selected interval. Consequently, over rough terrain the random method seems to be a better approach.

4. Results

4.1. Presentation of Results

Figures 2, 3, and 4 are illustrations, both of the signal received in the aircraft, and the terrain profiles of the paths over which the flights were made. The dashed lines indicate the range of maximums and minimums which would be expected over the smooth earth determined by the curve fitted to the terrain. The calculated smooth earth lobe structure to be expected along these paths is also presented. It is easily seen that the Fort Carson path is not one that can be adequately approximated by a simple curve; hence, no comparable calculated values are presented for this path. From this presentation several characteristics of received fields over rough terrain are evident. When both terminals were very high, the regular lobe structure characteristic of specular reflection did not exist. Nor was the extensive range of maximums to minimums apparent. Contrasted to this, the Fort Carson path with a low transmitting terminal exhibited both a definite lobe structure and a more extensive range of maximums to minimums. It may be well to question whether the restricted fading range was actually present or might perhaps be due to the limitations of the frequency response of the recording system. An investigation of the data and of the frequency characteristics of the recording system indicated that the recorded maximums and minimums at the near end of the paths; i.e., at the points labeled 32 miles from the transmitter, would not differ from the actual maximums and minimums by more than 25 percent, this possible error being due to recorder response limitations. This possible error diminishes rapidly and from the points 42 miles from the transmitter to the radio horizon all the data appear to be well within the response limitations of the recording system.

This singularity of the Fort Carson path can be interpreted as the presence of an extensive surface which would support specular reflection. The profile of the Fort Carson path does show that two distinct surfaces between the transmitter and a point 5 miles from the transmitter are relatively flat, and the geometry indicates that reflected energy from these regions is received over all points of the flight path. However, the geometry of the entire profile indicated that multiple reflection from as many as four general regions of the surface contribute to the field experienced at a point. This is consistent with the appearance of the received signal; i.e., that the specular reflection produces lobing, while random reflection produces the higher frequency fluctuations superimposed. Consequently, it can be seen that if one terminal is low, small specular reflecting surfaces can be effective in producing the characteristic lobing of specular reflection; but if the terminal is raised sufficiently such that the specular reflecting surface becomes a small part of the illuminated zone, these surface anomalies become less effective and specular reflection less apparent. Thus, it is inferred that reflection from rough terrain contains two components, a specular component and a random component. Where one terminal is low the specular contribution can be significant, but this contribution will decrease as the height of the lower terminal is increased. Therefore, it is interpreted that the data from two paths are illustrations of the sum of only two significant components, a constant vector and a Rayleigh-distributed vector, while the Fort Carson data, in addition, include a significant ground-reflected specular component.

Since two components in the ground reflection over the Fort Carson path can be identified, some method must be used to separate them. The difficulty of terrain approximation is such that no easy method can be devised based on path geometry. However, it seems reasonable to assume that the long-term spatial variations of the field are due only to specular reflection, and that the short-term spatial variations are the results of random reflection. If it can be assumed that the short-time rate-of-change of the effective amplitude due to the specular component is negligible in comparison to that due to the random component, then the medians of the instantaneous variations over short intervals accurately illustrate the relative field due to the specular component of the reflected waves plus the direct wave (fig. 5). Then the value of this specular component can be found using eq (1). It should be noted that this specular component is not necessarily a contribution from a single reflecting region, but is rather a summation resulting in a wave appearing as a single coherent wave at the receiver. Then, in this case, our "constant vector" is defined as consisting of a direct wave plus a coherent specular-reflected wave and to be slowly but uniformly changing due to the changing phase difference between the two coherent components. To compute the rms amplitude of the random reflection coefficient, a moving average can be used to eliminate the effect of the specular component and the analysis can then be done as outlined above.

Figures 6, 7, 8, show the distributions of the estimates of the rms amplitude of the random reflection coefficient as a function of grazing angle for each flight as determined by the above method. It was pointed out previously that this treatment assumes that the average energy reflected is constant over the interval chosen; i.e., k^2=constant.

The rms value of the random component was used only for convenience in that it was identically equal to k. The analysis would be approximately the same if mean values or median values were used, as

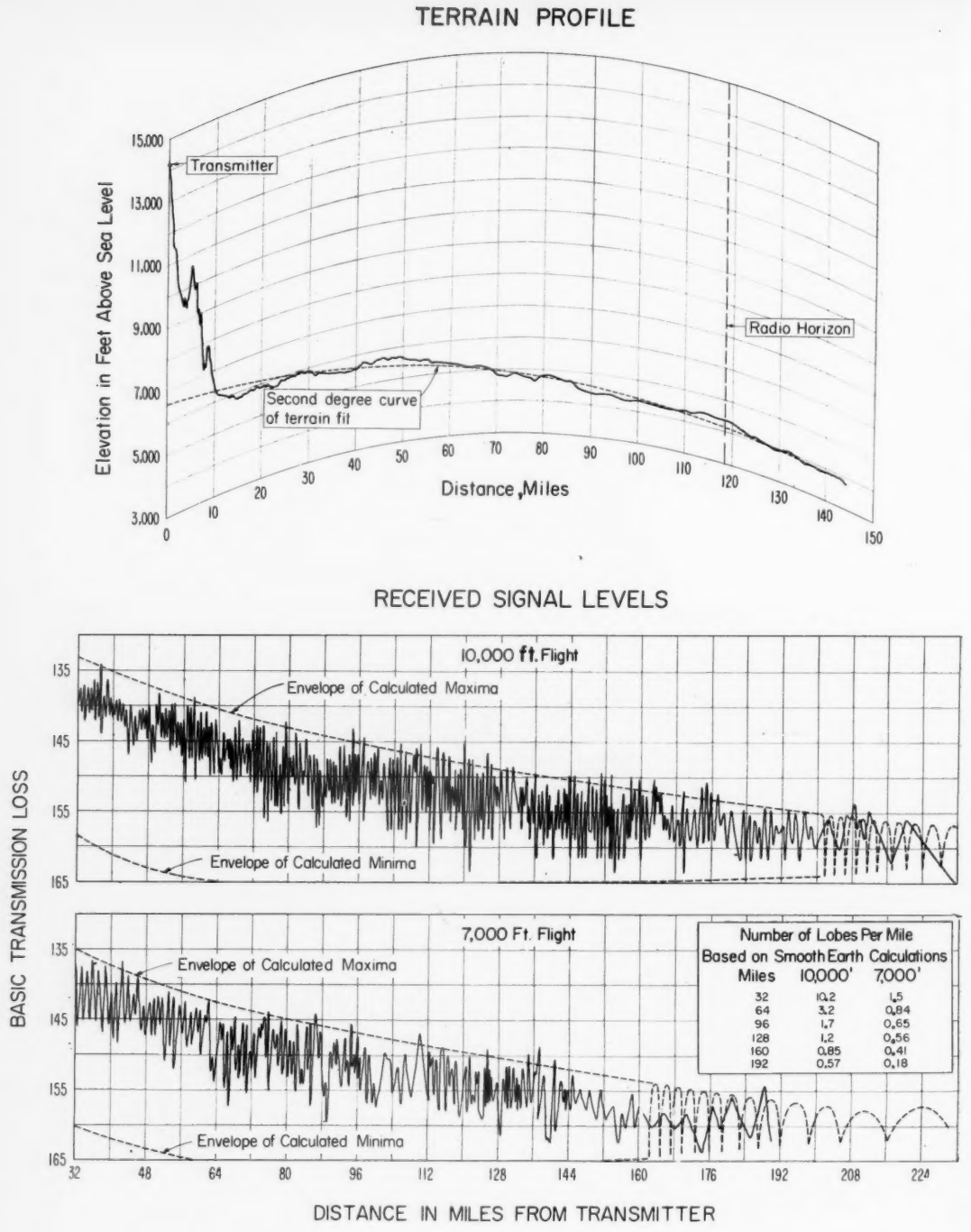


FIGURE 2. *Pikes Peak to Haswell path.*

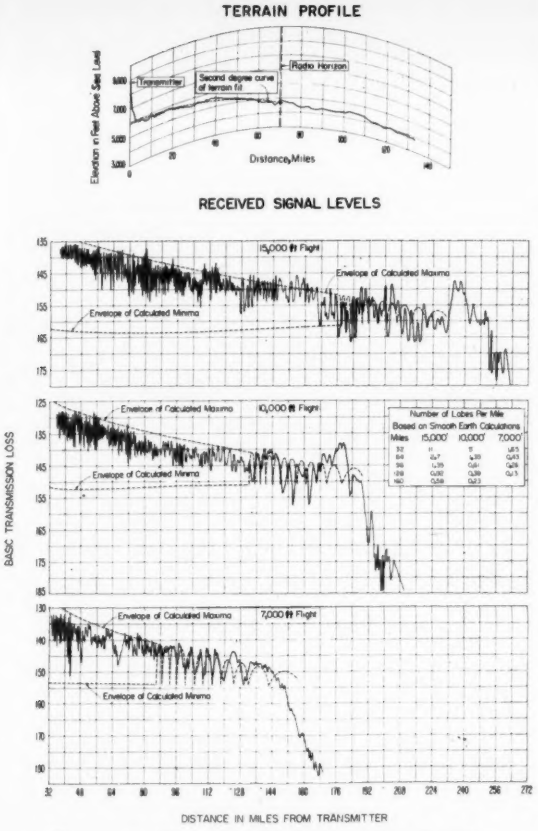


FIGURE 3. *Cheyenne Mountain to Haswell path.*

the Rayleigh distribution is completely specified when any one pertinent parameter is specified. For convenience here, the rms random component is to be interpreted as the rms value of the Rayleigh distribution. Since many considerations of Rayleigh distributions involve median values, it may be useful to note that the median value is equal to 0.8326 times the rms value.

4.2. Interpretation of the Results as a Function of Grazing Angle and/or as a Function of Terminal Height

It is reasonable to expect one or both of two possible relationships, the first, (a) that the rms random reflection coefficient is a function of the grazing angle, and/or the second, (b) that the rms random reflection coefficient is a function of the terminal height.

(a) To investigate the former, each flight was considered individually and regression analysis was performed to determine the trend. The regression coefficients of higher than the first degree terms were statistically insignificant; hence, linear regression was used. The linear regression lines are included on the appropriate figures 6 through 8. Table 2, pt. (), summarizes the statistical treatment of each individual flight as a function of the smooth earth grazing angle. Four of the seven correlation coefficients are statistically significant at the 5-percent probability level; however, two are positive and two are negative with no known explanation for the inconsistency. The significance may be unwarranted due to the fact that the observations may not be truly independent; whereas, the use of significance tests assumes independence. At any rate, the correlation coefficients are small, the largest being only 0.42, indicating, that, in the extreme, only 18 percent of the variations in the reflection coefficient can be attributed to variations in the grazing angle. Therefore, it is reasonable to conclude that the rms random reflection coefficient was only slightly dependent on the grazing angles for grazing angles less than 5°.

(b) The second possibility is that the rms random reflection coefficient was a function of terminal height. Since in the previous consideration it was found that there was little relationship between the grazing angle and the value of these reflection coefficients, the logical approach here would be to test the hypothesis that there is no difference among the results of the different flights. This involves an analysis of variance and covariance. The process is somewhat tedious and only the results will be presented here. The reader is referred to any standard text on statistical analysis [18, ch. 7]. Four of the seven flights appear to come from a single population having a single mean and single variance. This appears to contradict the conclusion of significance for the correlation coefficients, since two of these flights (the flights on the Pikes Peak path) were found to have significant slopes. The Fort Carson data differ from the four flights in table 2, part (2), not due to the means but due to the larger variances. The Cheyenne Mountain, 7,000-ft flight results differ from the remainder, due both to the low value of the mean and to the excessive slope. The departure of the Cheyenne Mountain 7,000-ft data cannot be adequately explained. If, for the sake of convenience, this single flight is ignored the remainder of the data can be interpreted. The reasonable conclusion would be that with the lower terminal 35 ft or higher, the mean value of the rms random component of the reflection coefficient does not vary with height in the range of heights investigated, but that increased variability about the mean was to be expected as this height approached 35 ft as a lower limit. Hence, it appears that for the conditions of this experiment the only effect of change of the lower terminal height was to alter the degree of variability in the instantaneous value of the rms random component of the reflection coefficient about its mean value.

4.3. Interpretation of the Results as Samples From a Single Population

The absence of a substantial correlation between the above mentioned parameters suggests further consideration regarding the scatter of the observed results.

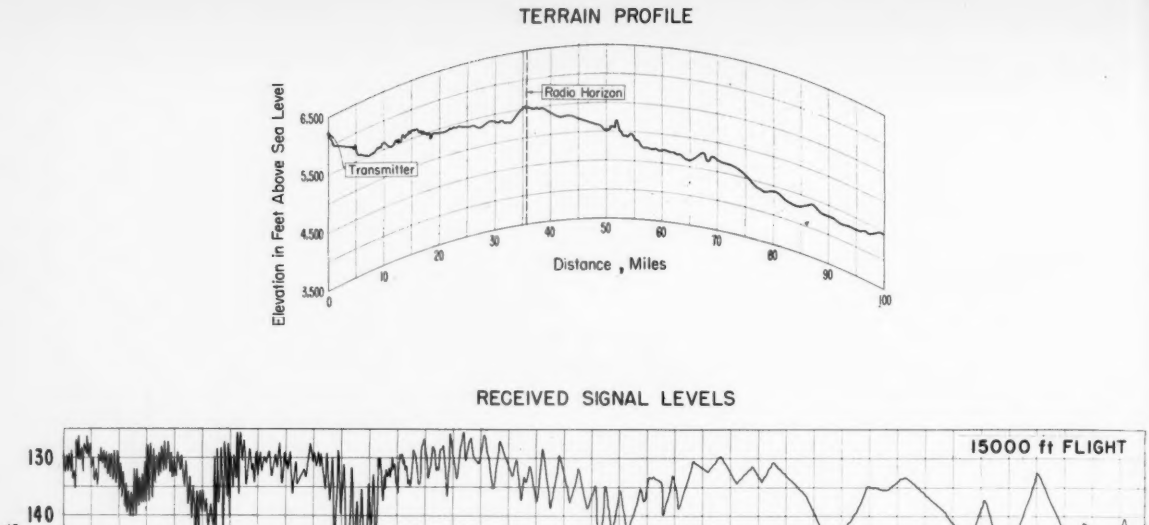


FIGURE 4. *Fort Carson to Haswell path.*

TABLE 2. *Statistical parameters for the random component of the reflection coefficient*

	No. of samples	Mean random component of the effective reflection coefficient	Standard deviation from mean	Linear regression line	Standard deviation from the regression line	Correlation coefficient
Part (1) Individual Flights:						
Fort Carson 10,000 ft	55	0.329	0.158	$\rho=0.213+0.112\psi$	0.145	[a] 0.418
Fort Carson 15,000 ft	75	.374	.239	$\rho=.437-0.041\psi$	.238	−.148
Cheyenne Mountain 7,000 ft	58	.261	.142	$\rho=.186+0.148\psi$	.132	[a] .387
Cheyenne Mountain 10,000 ft	65	.337	.120	$\rho=.307+0.039\psi$	.118	.207
Cheyenne Mountain 15,000 ft	136	.350	.140	$\rho=.365-0.013\psi$	.140	−.077
Pikes Peak 7,000 ft	90	.362	.108	$\rho=.413-0.029\psi$	.101	[a] −.319
Pikes Peak 10,000 ft	117	.391	.130	$\rho=.441-0.034\psi$	.127	[a] −.215
Part (2) Flights which indicate statistically homogeneous data:						
Cheyenne Mountain 10,000 ft + 15,000 ft; Pikes Peak 7,000 ft + 10,000 ft	408	.362	.128			
Part (3) Total Data						
All flights	596	.351	.149			

[a] Statistically significant at the 5% probability level.

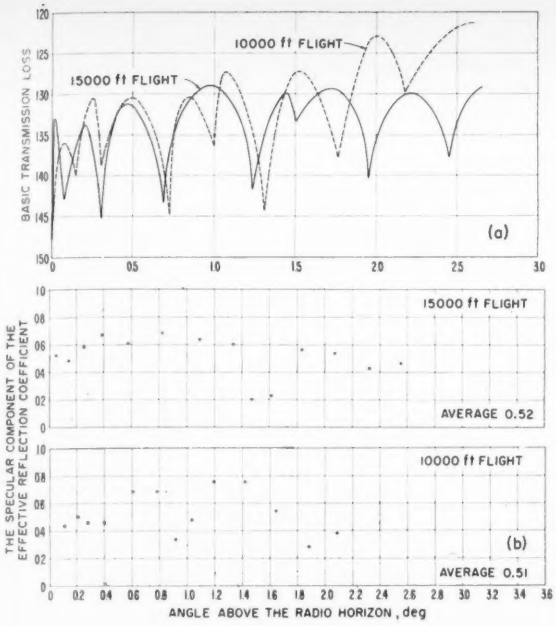


FIGURE 5. *Fort Carson path.*

a. Distribution of received signal showing only the long-term spatial variations attributed to a specular component. b. Distribution of the specular component of the effective reflection coefficient.

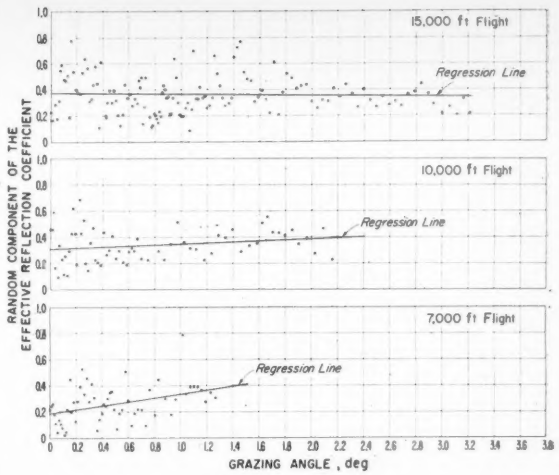


FIGURE 7. *Cheyenne Mountain path.*

Distribution of the effective reflection coefficient.

It should be remembered that each interval was analyzed as if it were an independent sample which is probably not strictly true, and that each "value" of the random reflection coefficient is an estimate of the rms value of a Rayleigh-distributed variate for that interval. Consequently, the results presented in figures 6, 7, and 8 are distributions of estimates of the rms values of a series of approximately independent Rayleigh distributions. It is reasonable at this point to consider the possibility that these estimates of the rms value of all of these independent Rayleigh distributions may be normally distributed. This would be equivalent to saying that the form of the cumulative distribution would remain the same, but that the value of the intercept would be normally distributed. The method of analysis indicates a possibility that the estimates of any of three parameters may be approximately normally distributed. The parameter K, where $K=10 \log k^2$, could indicate a log-normal relationship; the reflected energy (proportional to k^2) may be normally distributed; or the rms random reflection coefficient (equal to k) may be normally distributed. Kolmogorov's test, as applied by Barrows and Norton [19], was used to test the normality of the distribu-

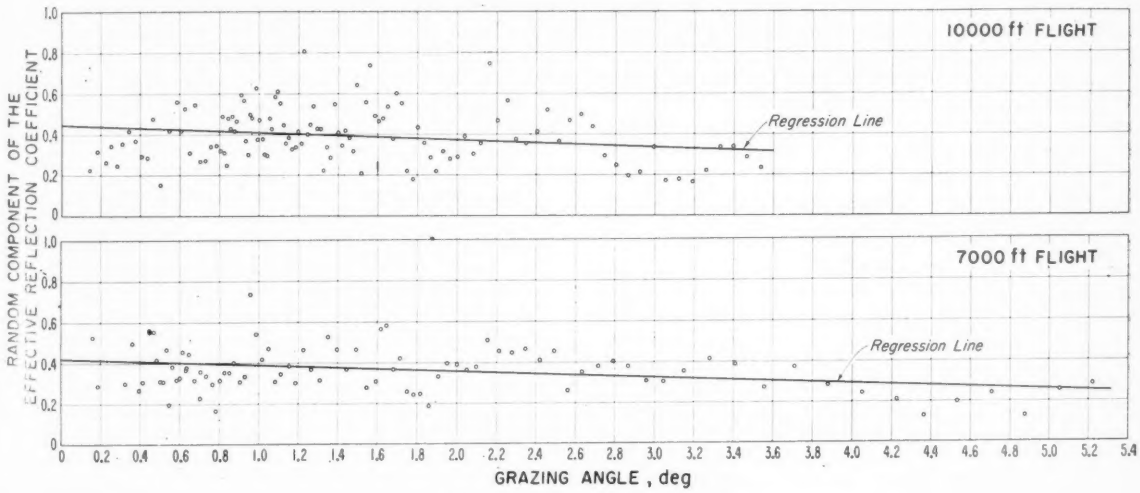


FIGURE 6. *Pikes Peak path.*

Distribution of the effective reflection coefficient.

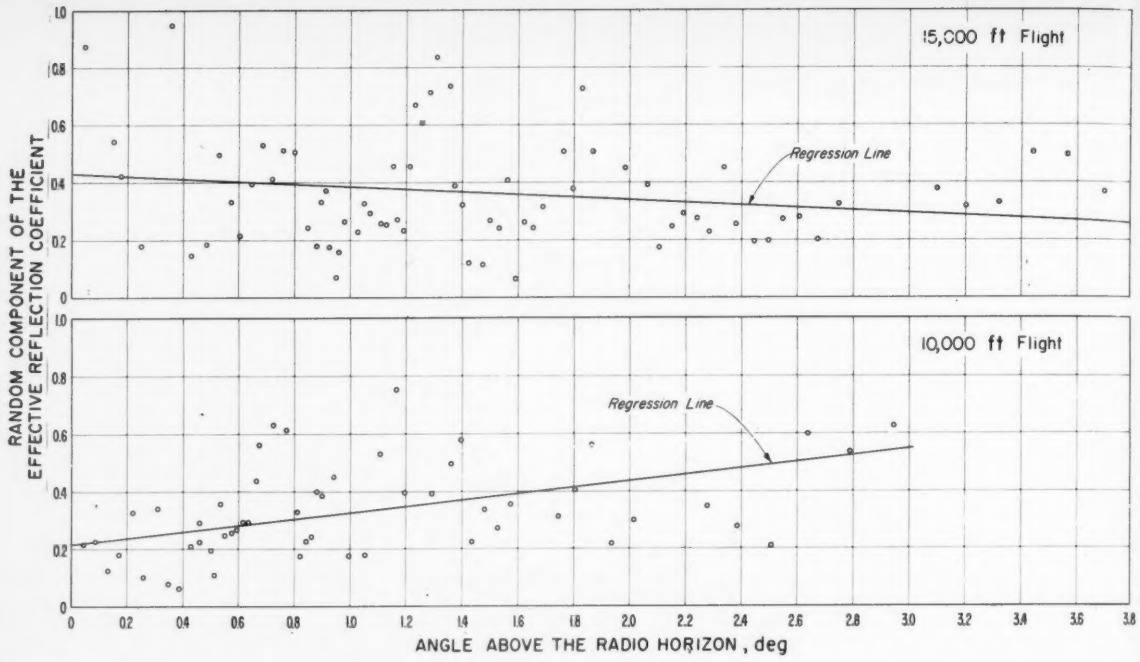


FIGURE 8. *Fort Carson path.*

Distribution of random component of the effective reflection coefficient based on short-term spatial variations of received signal.

tions of these three estimates. This is not an exact test, but a reliable approximation. The test was used in both directions; i.e., (1) on each flight as a separate unit, and (2) on each interval of grazing angle utilizing the data from all flights. As a result of testing each flight as a separate unit, and testing each interval of grazing angle, the normal distribution was found to approximate the distribution of k better than the distributions of k^2 and K. When the data were considered as a whole, all three parameters appeared normally distributed, however, k again proved to be the best approximation.

It was pointed out earlier that statistically these flights were not samples from the same population; consequently, when all of the data are considered as a unit, some error will be introduced. There are apparently three populations of data: the Cheyenne Mountain 7,000-ft data; the Fort Carson data; and the remainder which form a homogeneous group. The amount of error introduced by such an amassing should be relatively small. The low-correlation coefficients, even though significant, indicate that little error is introduced when the results of each flight are considered as samples of a distribution having a single mean. Considering the results of all flights as members of a single population requires only that it be remembered that the lower antennas introduce a higher variability from the mean value. (The departure of the Cheyenne Mountain 7,000-ft results is not explained.) Comparing the results of considering the data as a whole, with only that part of the data indicated to be homogeneous, little difference is noted. In the former case, the mean is 0.35 with a standard deviation of 0.15 for 596 observations; in the latter case, the mean is 0.36 with a standard deviation of 0.13 for 408 observations. Comparison of the amassed results can also be made up with the inhomogeneous group; i. e., the Fort Carson data where the 15,000-ft flight shows the greatest departure. The mean of this flight is 0.37 and the standard deviation is 0.24, which illustrates the tendency for greater variability when one terminal is very low.

If this conclusion is utilized in two steps, further evidence of a lack of significance is seen. Considering first, that the random reflection coefficient is independent of terminal height, and considering that estimates of the rms reflection coefficient would be normally distributed within small intervals of the grazing angle, it is possible to reevaluate the random reflection coefficient as a function of the grazing angle. The mean and standard deviation was found for each interval of the grazing angle utilizing the estimates of k from all seven flights. The result of this procedure are shown in figure 9. The 90-percent confidence limits are only approximate since the data are not from a truly single population. The point at 1.3° seems to be excessively high. This is due chiefly to the contributions from the Fort Carson data at this point, and can be interpreted as caused by possible anomalies in the Fort Carson path. It is in this region that the assumption of a single population is poor. It appears that the distribution of the Fort Carson

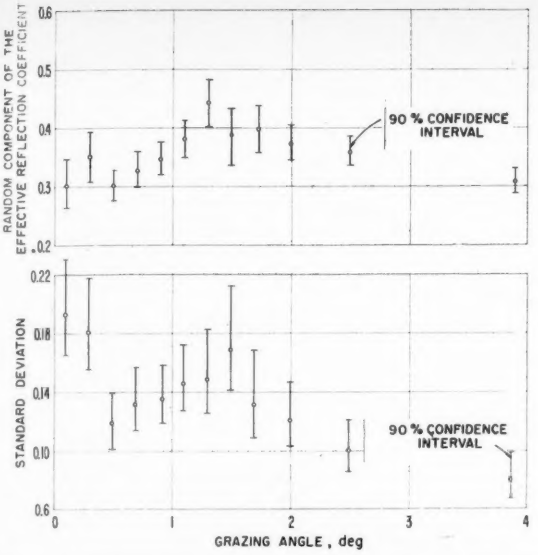

RANDOM COMPONENT OF THE EFFECTIVE REFLECTION COEFFICIENT
90 % CONFIDENCE INTERVAL
STANDARD DEVIATION
90 % CONFIDENCE INTERVAL
GRAZING ANGLE, deg

FIGURE 9. *Random component of the effective reflection coefficient as a function of the grazing angle for all flights.*

data may be the sum of two normal distributions having different means; a departure such as this can be caused by the addition of a small number of nonidentical populations. Although this is indicated, it is not good statistical procedure to dissect data unless more justifiable reasons are evident [20].

As a second consideration, the reflection coefficient can be assumed to be independent of the grazing angle, and the estimates of the rms value of the random reflection coefficient can be considered to be normally distributed for discrete terminal heights. Here, the mean and standard deviation, as shown in table 2 for each flight, can be expressed as a function of transmitter height. Such results would reevaluate the random reflection coefficient as a function of terminal height. Figure 10 is an illustration of such a consideration. With the exception of the Cheyenne Mountain 7,000-ft flight indicated at a transmitter height of 2,836 ft and a receiver height of 1,625 ft, there is little difference in the value of the random component of the reflection coefficient. It should be noted that the standard deviation decreased with increase of terminal height.

Kolmogorov's test applied to the total data confirmed the assumption of normality. Figure 11 illustrates the cumulative distribution of the total data. The mean for the total data is 0.35 and the standard deviation is 0.15 (from table 2), which compares well with the median value of 0.33 and a standard deviation of 0.15 (from fig. 11).

The proper interpretation of these results is that the rms random component here stated is the rms value of a Rayleigh distribution, and that the estimate of this rms value is approximately normally distributed having a mean value of 0.35 with a standard deviation of 0.15. If the Rayleigh distribution were expressed graphically on log-Rayleigh coordinates, the slope would remain -1, but the intercept at Pr $(\rho > k) = 0.368$ would be approximately normally distributed about 0.35 (fig. 12).

No doubt the standard deviation is a function of the length of the sampling interval as well as being a function of the different populations amassed

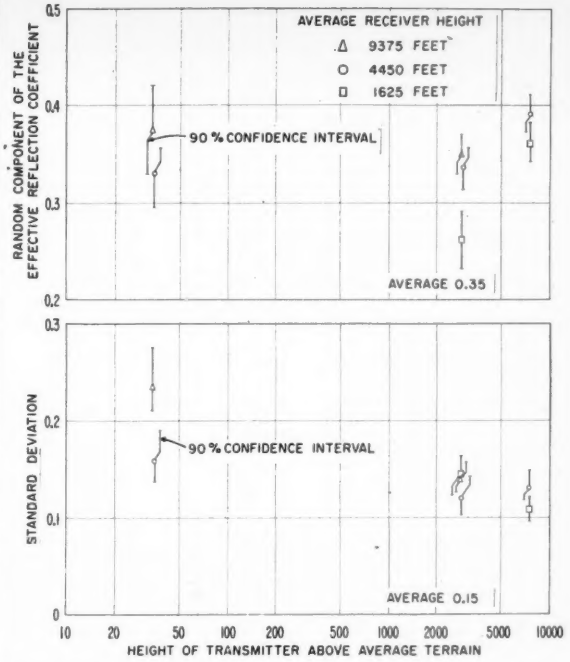


FIGURE 10. *Random component of the effective reflection coefficient as a function of transmitter height for all flights.*

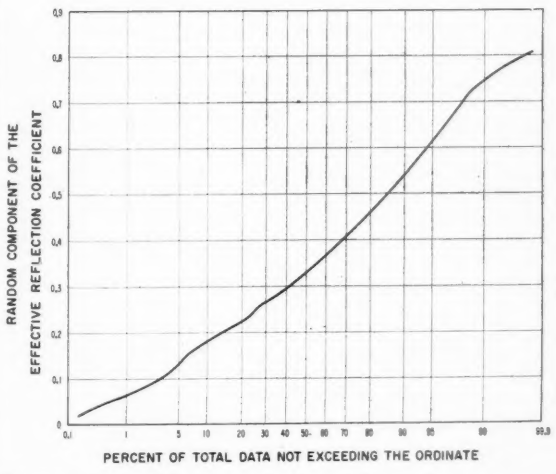


FIGURE 11. *Cumulative distribution of the random component of the effective reflection coefficient.*

Total data for all flights—596 points.

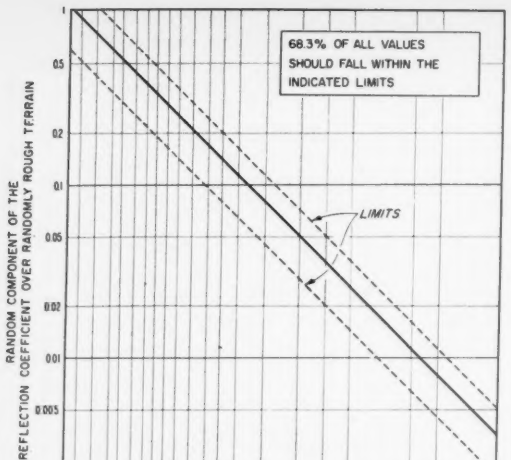


FIGURE 12. *Distribution of random component of reflection coefficient for grazing angles less than 5°.*

It is expected that 68.3 percent of individual values should be found to be between the indicated limits with the probability expressed along the abscissa.

in the total data. Thus it is difficult to interpret, i.e., it is composed of two parts; one is the variation of sample estimates of a population parameter, the rms random reflection coefficient; the other is the variation due to nonhomogeneous populations. No attempt is made here to indicate how much of this standard deviation is due to each source of variation.

It should be emphasized that this value for the random reflection coefficient ignores any divergency caused by reflection from the curved surface of the earth.

5. Comparison With the Results of Other Investigations

Due to the inability of accurately determining the reflecting surface, it was not possible to evaluate quantitatively the relationships expressed by some authors in terms of the type of analysis made here. However, in those cases where comparison was possible, remarkably good agreement was found particularly where overland measurements were considered.

Bullington's [21] data on the New York to Denver path compares favorably. Assuming a Rayleigh distribution for Bullington's data as suggested by Rice [5], we find a median value of 0.29 and an rms value of 0.35. The variance is roughly the same. Straiton [22] utilizing four different frequencies over both land and water, achieved similar results for the higher frequencies over land. Similarity is found with the results of Sherwood and Ginzton [23] in the case of rough terrain. Although a proper comparison cannot be made with the results published by Beard, Katz, and Spetner [24], it is significant that comparable conclusions concerning the constancy of the random component with increasing effective roughness is also proposed by these authors. This constant mean value is quickly achieved for small values of the parameter expressing effective surface roughness. It should be noted that the results of these different investigators were found in widely separated places where ground constants can be assumed to differ. It would appear that the results are independent of the value of ground constants providing an overland path is considered.

In most cases where paths over rough water are considered, the value of the reflection coefficient for low grazing angles appears to be much higher and to approach more closely the smooth earth expected value [7, 22, 23, 24].

6. Conclusion

Thus, we arrive at the following description. The reflected signal received from rough terrain even at very small grazing angles can be considered to be made up of two components, one that is specular, and the other being Rayleigh distributed. Assuming one terminal as the reference, maintained above 1,000 ft, the relative magnitude of the two components of the reflection coefficient can be viewed as a function of the height of the variable terminal. Where the variable terminal is very low, the Rayleigh component may be considered to be small with respect to the specular component, but to increase in relative magnitude as the height of the low terminal is increased. Although in this experiment the rms random component exhibited essentially the same mean value for varying terminal heights, it would be expected that the value of the rms random component would approach zero as the height of the terminal approached zero [23]. The variance appears to decrease as the lower terminal-height increases especially at very low terminal heights. Therefore, the variability of the random component would be expected to be greater with a low terminal, diminishing as this terminal is raised. The overall mean value of the random component of the reflection coefficient was found to be 0.35; the standard deviation of individual estimates of the rms random component obtained from distance intervals of the flight paths as outlined above, was found to be 0.15.

Although horizontal polarization was used in this investigation, S. O. Rice [5] points out that over rough terrain the identical mechanism is operative regardless of polarization. This view seems to be borne out in the overland data of Straiton [22] and Bianco and Morris [25], both of whom notice no polarization effect if the terrain is sufficiently rough. Thus, it is reasonable to assume that similar results would have been obtained along these paths if vertical polarization had also been investigated.

The results presented here viewed in comparison with the reports of others suggest further considera-

tion of the problem of rough terrain reflection. It appears that there are many variables present which are not wholly understood. At present, it would appear that, over a spherical earth, if the terrain is sufficiently rough and sufficiently extensive, the value of the random reflection coefficient might be essentially independent of frequency, grazing angle, polarization, and ground constant; having a mean value of approximately 0.3.

The variability of the instantaneous value is rather large and is more pronounced for low terminals. Specular reflection can be expected over rough terrain particularly if the terminals are low. There appears to be a difference between sea and land reflection in so far as rough terrain reflection is concerned, but at present it is not possible to determine whether this is due to marked difference in ground constants or to the degree of effective roughness.

It should be obvious that some liberty was taken with mathematical exactitude in the analysis of the data. This investigation points out what relationships appear to be present, but it should be apparent that additional study is required. Although it might appear that considerable data were present for this study, it is evident that a full investigation of this particular phenomenon will require considerably more than was available for this presentation. Such studies could lead to a possible macroscale factor for specifying rough terrain as a function both of frequency and antenna height.

7. Addendum, An Estimate of the Relationship Between the Fading Range and Terminal Heights

Since it is apparent that specular reflection was more effective than random reflection in providing large spatial fading ranges, it would be advantageous to determine a criterion for elimination of the specular component. Since it seems that low terminals produce a specular component even over rough terrain whereas high terminals produce only random reflection, an investigation of the relationship between terminal height and the prevalence of the specular component would be desirable. It is felt that the present data are not sufficiently complete to provide more than an indication. The spatial fading range for the Fort Carson data is interpreted to be 16.5 db on the average. Since there has not been assumed geometrical relationship for the shape of the terrain, even the specular component is not corrected for divergence and the values expressed are effective values at these points. Regression analysis indicates that there was no significant correlation between the value of the effective specular reflection coefficient and the grazing angle. Hence, in computing the spatial fading range, the mean value of the specular component for the entire range (0.51) was used. The average maximum and minimum values of the sum of a direct wave and specularly-reflected wave whose coefficient is 0.51 were determined. Combination of these values with a random component having a proper rms value as expressed in table 2 yielded fading ranges of 16 and 17 db. (The spatial fading range as previously defined is the ratio of the fields exceeded 10% and 90% of short intervals of distance along the path.) Assuming that the higher terminals produce only random reflection with reflection coefficients as expressed in table 2, spatial fading ranges of less than 7 db were encountered. Figure 13 is a preliminary estimate of the relationship between the spatial fading range and the height of the lower of two terminals (maintaining the other terminal above 1,000 ft) for grazing angles less than 5°. The unlabeled point is the average value of the spatial fading range for the 7,000-ft flights. It would appear that the minimum spatial fading range to be expected under optimum conditions would be in excess of 5 db. If a fading range of 6 db could be tolerated, then the minimum lower antenna height would be 600 ft, whereas 270 ft would be sufficient for a 7-db fading range. Of course, these relationships are based on the validity of the assumption of the shape of the curve of figure 13. Many other curves including a straight line could be assumed with equal mathematical justification. The authors present this estimate as a reasonable approximation consistent with the data considered. Additional investigation is necessary before such an estimate could be accepted as typical of rough terrain phenomenon.

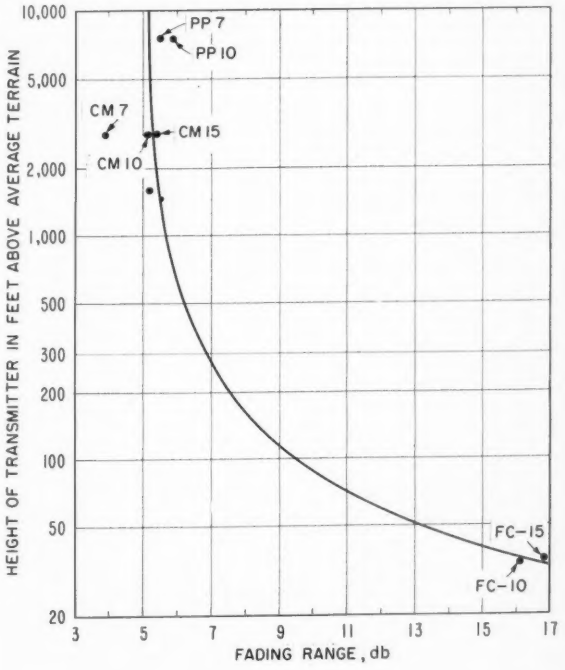


FIGURE 13. *Estimate of the fading range as a function of terminal height over randomly rough terrain at 1,046 Mc for grazing angles less than 5°.*

Receiver or reference terminal in excess of 1,000 ft.

The authors acknowledge the cooperation of the Wright Development Center of the Wright-Patterson Air Force Base in supplying the B–17 aircraft. In addition, the authors acknowledge the assistance of Kenneth A. Norton, Jack W. Herbstreit, Kenneth O. Hornberg, and Dr. Edwin L. Crow of NBS, both in the presentation and in the review of this paper.

8. References

[1] R. S. Kirby, J. W. Herbstreit, and K. A. Norton, Service range for air-to-ground and air-to-air communications at frequencies above 50 megacycles, Proc. IRE **40,** 535 (1952).

[2] K. A. Norton and A. C. Omberg, The maximum range of a radar set, Proc. IRE **35,** 4 (1947).

[3] H. Davies, The reflection of electromagnetic waves from a rough surface, Proc. Inst. Elec. Engrs. (London) **101, Pt. IV,** 209 (1954).

[4] V. Twersky, On the non-specular reflection of electromagnetic waves, J. Appl. Phy. **22,** 825 (1951).

[5] S. O. Rice, Reflection of electromagnetic waves from slightly rough surfaces, Communs. Pure and Appl. Math. **4,** 351 (1951).

[6] W. S. Ament, Toward a theory of reflection by a rough surface, Proc. IRE **41,** 142 (1953).

[7] Peter Beckman, A new approach to the problem of reflection from a rough surface, Acta Tech. Ceskosl. Akad. Ved **2,** 311 (1957).

[8] A. I. Kalinin, Approximate methods of calculating the field strength of ultra-short radio waves, taking into consideration the influence of the local terrain, Radio Tech. (Transl. Pergamon Inst.) **12,** 13 (1957).

[9] M. A. Biot, Some new aspects of the reflection of electromagnetic waves on a rough surface, J. Appl. Phys. **28,** 1455 (1957); Reflection on a rough surface from an acoustic point source, J. Acoust. Soc. Am. **29,** 1193 (1957).

[10] J. R. Wait, Guiding of electromagnetic waves on a uniformly rough surface: parts I and II, Proc. International Symposium on Electromagnetic Theory, University of Toronto, June 15–20, 1959.

[11] A. P. Barsis, J. W. Herbstreit, and K. O. Hornberg, Cheyenne Mountain tropospheric propagation experiments, NBS Circ. 554, (1955).

[12] K. A. Norton, Transmission loss of space wave propagation over irregular terrain, IRE Trans. **PGAP–3,** 152 (1952).

[13] D. E. Kerr, Propagation of short radio waves, p. 396 (McGraw-Hill Book Co., New York, N.Y., 1951).

[14] C. R. Burrows and S. S. Attwood, Radio wave propagation, p. 466 (Academic Press, New York, N.Y., 1949).

[15] K. A. Norton, Propagation over rough terrain, U.S. NEL Symposium on Tropospheric Wave Propagation, July 25 to 29, 1949, Naval Electronics Rept. No. 173, p. 101.

[16] K. A. Norton, L. E. Vogler, W. V. Mansfield, and P. J. Short, The probability distribution of the amplitude of a constant vector plus a Rayleigh distributed vector, Proc. IRE **43,** 1354 (1955).

[17] S. O. Rice, Mathematical analysis of random noise, Bell System Tech. J. **23,** 282 (1944) and **24,** 46 (1954).

[18] C. A. Bennett and N. L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry (John Wiley & Sons, New York, N.Y., 1954).

[19] E. Barrows and K. A. Norton, The Kolmogorov test of the goodness of fit of data samples to independently specified continuous distributions, together with an exact test of the normality of a small sample (to be published).

[20] A. Hald, Statistical theory with engineering applications, p. 152 (John Wiley & Sons, New York, N.Y., 1952).

[21] K. Bullington, Reflection coefficients of irregular terrain, Proc. IRE **42,** 1258 (1954).

[22] A. W. Straiton, Microwave radio reflection from ground and water surfaces, IRE Trans. **PGAP–4,** 37 (1952).

[23] E. M. Sherwood and W. L. Ginzton, Reflection coefficients of irregular terrain at 10 cm, Proc. IRE **43,** 877 (1955).

[24] C. I. Beard, I. Katz, and L. M. Spetner, Phenomenological vector model of microwave reflection from the ocean, IRE Trans. **AP–4,** 162 (1956).

[25] D. R. Bianco and C. S. Morris, Radar terrain cross section at microwave frequencies, delivered at the fall Joint Meeting of URSI–IRE, Penn. State Univ., Univ. Park, Penn., (Oct. 20 to 22, 1958).

Boulder, Colo. (Paper 63D2–20.)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Synoptic Study of the Vertical Distribution of the Radio Refractive Index[1]

B. R. Bean, L. P. Riggs, and J. D. Horn

(March 12, 1959)

An analysis of the vertical structure of an intense outbreak of continental polar air is presented in terms of the radio refractive index of the atmosphere. Employed for the first time is a reduced index analogous to potential temperature. The reduced value more clearly shows the refractive index structure than the classical methods used heretofore. This new unit is a measure of both atmospheric density and humidity and shows, on a single cross section, the airmass structure and the dynamic mixing of air around the frontal interface.

1. Introduction

The problem of determining the vertical and horizontal distribution of the index of refraction of the troposphere for radiofrequencies has engaged the attention of radio meteorologists for the past decade [1, 2, 3, 4][2]. Such information is vital to procedures and techniques for forecasting the occurrence of radio refractive-index profiles. The surface value and the gradient of the refractive index are important factors in the prediction of monthly median values of field strength over communication links utilizing frequencies from 50 to 50,000 Mc [5, 6, 7].

For convenience, the refractive index is given in terms of its scaled up value, N, commonly called the refractivity:

$$N=(n-1)\ 10^6.$$

N may be determined [8] from standard weather observations by

$$N=(n-1)10^6=\frac{77.6}{T}\left[P+4810\,\frac{e_s RH}{T}\right],$$

where

n=index of refraction of the atmosphere,
T=temperature in degrees Kelvin,
P=station pressure in millibars,
e_s=saturation vapor pressure, and
RH=relative humidity, in percent.

2. Background

A worldwide survey of N climatology [9] indicates that the refractive index varies widely over the surface of the earth. If we take the modern view that climatology is a synthesis of synoptic events then we may conclude that the frequency of occurrence of airmass types and frontal passages play an important role in determining the refractive-index climate at any location.

Recent studies have shown that the N profile at a given location varies with synoptic pattern [10] and that, if one reduces N to sea level, the refractive index is a sensitive indicator of large-scale weather systems [11] when considered on a daily weather map basis. The present study extends the work of reference [11] by considering if the air mass properties associated with a typical wintertime outbreak of polar air are reflected in the refractive-index structure. A storm sequence with strong contrasts of airmass properties was chosen for study under the premise that if the refractive index failed to indicate the differences in airmass properties in this case then there is little hope of its being effective in situations with less marked temperature and humidity contrasts.

Before proceeding one might inquire if the particular case chosen for study was a representative one or not. As a guide in this regard, idealized models (figs. 1 and 2) of cold and warm fronts were con-

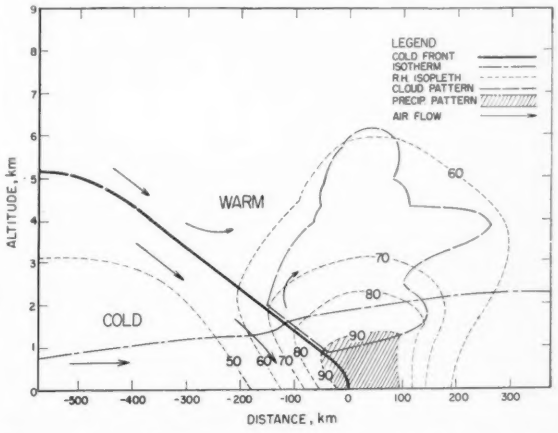


FIGURE 1. *An idealized cold front with associated temperature and humidity patterns.*

[1] This work was supported in part by Task 31 of the Navy Weather Research Facility, Naval Air Station, Norfolk, Va.
[2] Figures in brackets indicate the literature references at the end of this paper.

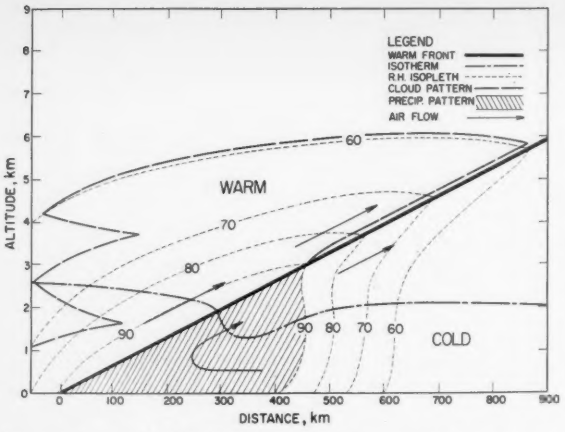


FIGURE 2. *An idealized warm front with associated temperature and humidity patterns.*

structed from generally accepted concepts of the polar front [12,13,14]. From these models, idealized N fields were constructed. This idealized N structure (figs. 3 and 4) is in agreement with that derived from a classic storm reported by Byers [13]. The most prominent feature of this idealized situation is the laminar structure of the N field reflecting a uniform decrease of N with height so pronounced that it masks the airmass contrasts about the front which are the main concern of this study. In the sections that follow we shall be concerned with the removal of the normal N decrease with height in order to more clearly see the N changes associated with the storm under study.

3. Height Dependence of N

It has long been recognized that the pressure term in N results in a steady decrease of N with increasing height. Early attempts to compensate for this N decrease used the constant gradient of the effective earth's radius theory, $1/4a$, where a is the radius of the earth. As an illustration, the strong elevated layer

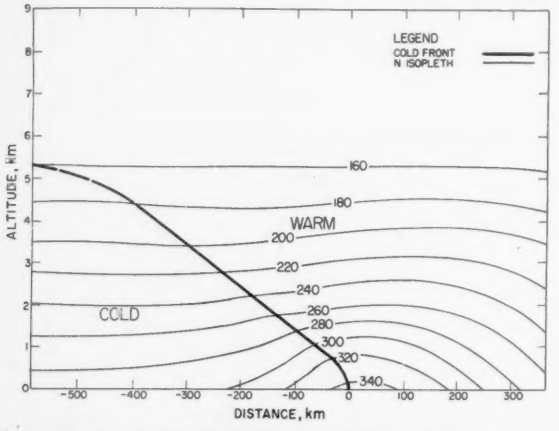


FIGURE 3. *The N field of an idealized cold front.*

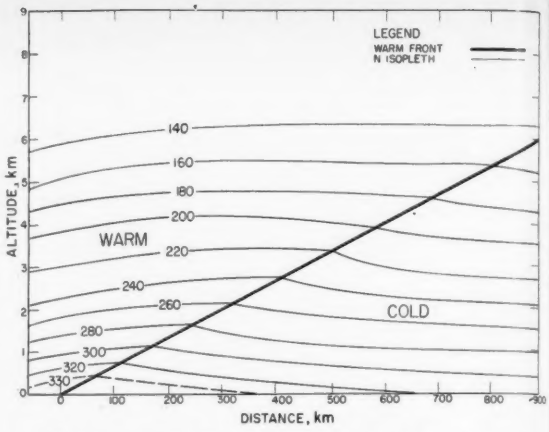


FIGURE 4. *The N field of an idealized warm front.*

found during the summer in Southern California was studied in terms of a linearly corrected expression, B, given by

$$B=N(h)+39.2h$$

where $N(h)$ is the value of N at any height, h, in kilometers [15]. Recent studies of the average N structure of the atmosphere show that N tends to be an exponential function of height rather than the linear function assumed by the effective earth's radius theory, with the result that the B unit approach overcorrects when h is greater than about 1 km. This is illustrated by figures 5 and 6 where the N data of figures 3 and 4 are replotted in terms of B units. Note that this over correction produces an N field that tends to increase with height from 310 at the surface to 360 at 5 km. If, however, one presents these data in terms of an exponential function, A, of the form

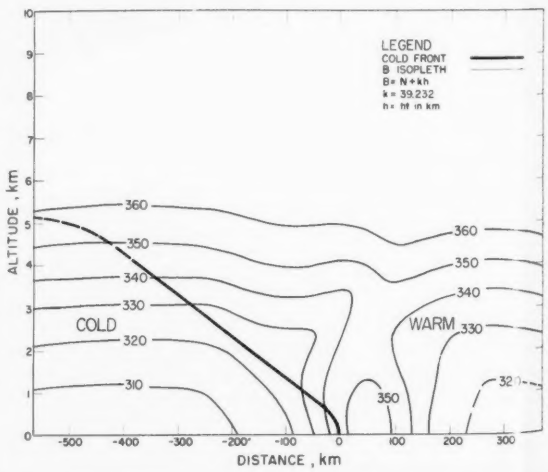


FIGURE 5. *The refractive index structure in* B *units for an idealized cold front.*

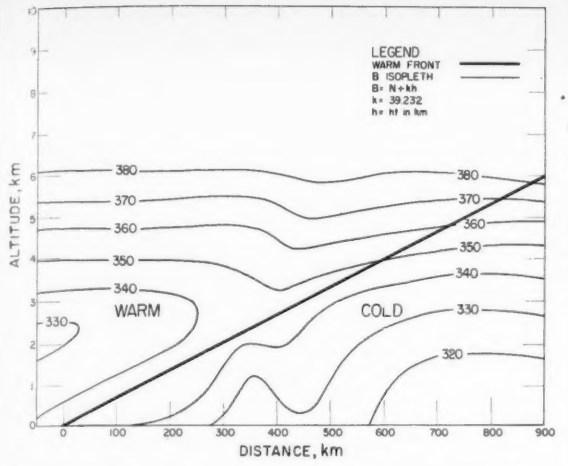


FIGURE 6. *The refractive index structure in* B *units for an idealized warm front.*

$$A=N(h)+313[1-\exp\{-0.144h\}],$$

as in figures 7 and 8, then it is seen that the range of N values is now reduced to about 25 N units. The exponential decay coefficient of 0.144 was determined to correspond with the average station value of N, 313 for the United States [16]. It is significant, however, that the use of A units produces a pattern similar to the airmass differences associated with the frontal zone. Note, in the warm front case (fig. 8), that the A values increase with height until they reach a maximum associated with the upgliding warm moist air overriding the frontal surface. The region of maximum precipitation 400 to 500 km in advance of the front is shown as an area of high surface N. In the cold-front case (fig. 7) the classic push of warm air aloft by the encroaching cold air is evidenced by the "dome" of high A values just before the front. Stratification in the cold air due to subsidence-inversion effects, although impossible to detect in the N charts, is clearly seen by the use of the A charts.

A units were used in all subsequent analysis in order to throw the frontal discontinuities and airmass differences into sharp relief. To facilitate this transformation, a Potential Refractivity Chart (fig. 9)

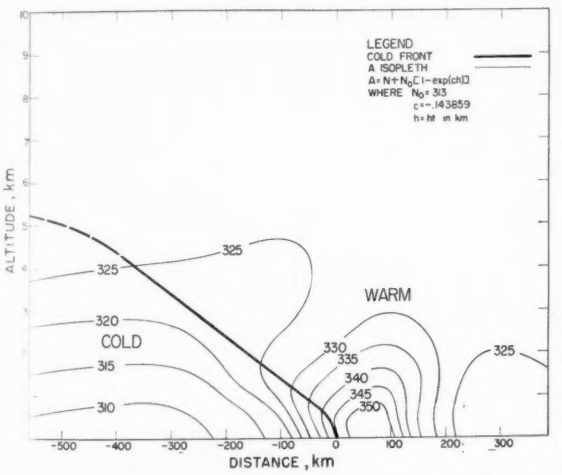


FIGURE 7. *The refractive index pattern in* A *units for an idealized cold front.*

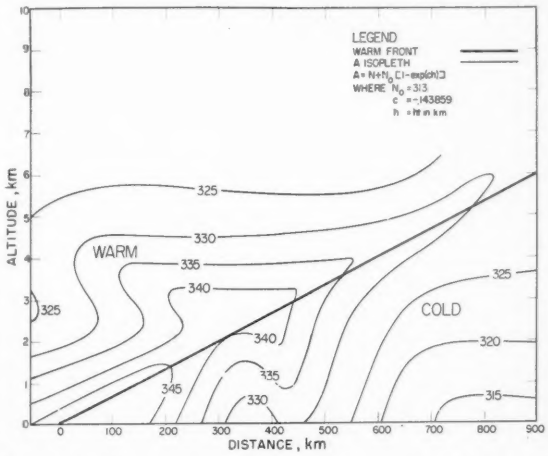


FIGURE 8. *The refractive index pattern in* A *units for an idealized warm front.*

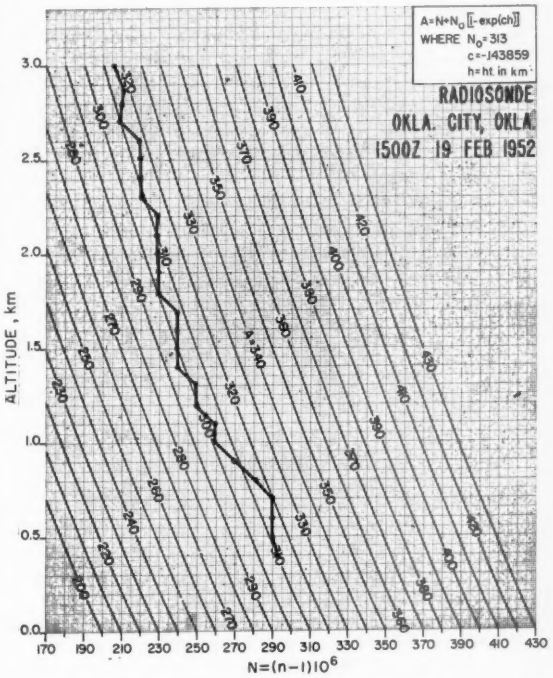


FIGURE 9. *Potential refractivity chart with radiosonde data shown in* A *units.*

A is obtained by plotting N versus altitude and interpolating from the lines of constant potential refractivity.

was prepared for rapid conversion from N units to A units. This simplification eliminates the necess ty of using exponential tables for each individual al-culation of A and thus lends considerable ease to the preparation of charts of the new parameter.

Storm of February 18 to 21, 1952

A large-scale outbreak of polar continental air which took place over the United States between the 18th and 21st of February 1952, was analyzed in terms of A units. Charts showing the structure of this outbreak of polar continental air have been prepared from the radiosonde measurements from stations located along a line normal to the front running from Glasgow, Mont. to Lake Charles, La. (fig. 10). The height distributions of unit contours along this cross-section line were obtained at 12-hr intervals during the 4-day period. Examples are given (figs. 11 to 13) to represent the early, mature, and late stages of the outbreak. At the outset of the period of observation, the A unit gradient was quite flat around the polar front. By 1500 Z on the 18th of February (fig. 11) the polar front was located midway between Glasgow and Dodge City. The contrast of the southward push of polar air and the northerly advection of tropical maritime air from the Gulf of Mexico into the developing warm sector of the polar-front wave is evidenced by the relatively large gradients in the neighborhood of Dodge City. The core of tropical maritime air has evidently not progressed far enough northward to displace the warm but dry air that had been over the great plains prior to the outbreak with the result that a region of low A values is found between the front and the tropical maritime air. Twenty-four hours later (fig. 12) the core of tropical maritime air has become more extensive and now reaches to 3 km. A second cold front was reported on the daily weather map and the area of low A values is now confined between

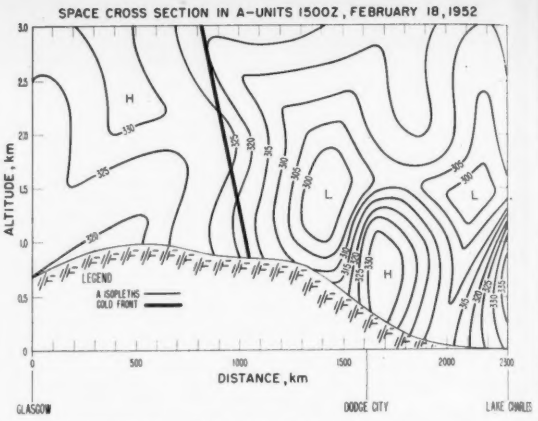


FIGURE 11. *Early stage of the storm sequence.*

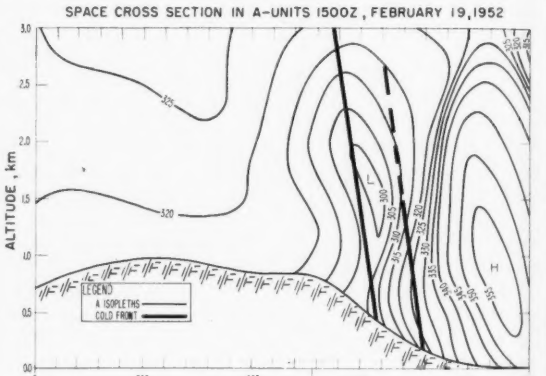


FIGURE 12. *Mature stage of the storm sequence.*

FIGURE 10. *Location chart of the stations furnishing radiosonde data for this study.*

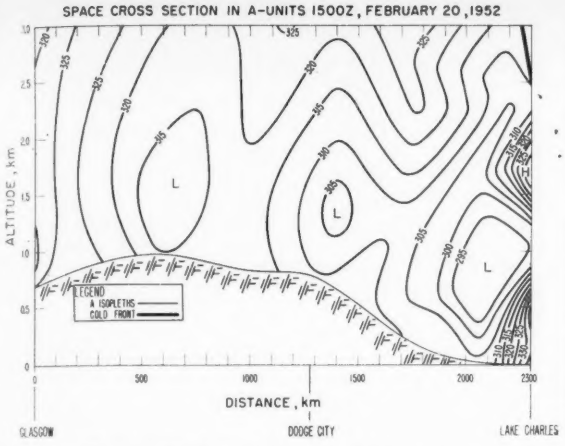


FIGURE 13. *Late stage of the storm sequence.*

it and the main cold front. Finally, by the morning of February 20 (fig. 13), the front has passed to the south of Lake Charles and the polar air just behind the front is characterized by relatively low A values.

The use of space cross sections does not always yield measurements at the most desirable points along a frontal zone. Another method of arriving at the probable refractive-index structure about the frontal interface is to present each successive radiosonde observation at a single location in terms of the distance between the front and the station. Thus, as the frontal system advances and passes over the station, one obtains a different perspective of the space cross section. This presentation is referred to as an epoch chart since the observations are normalized with respect to the frontal passage. Such a presentation is given on figures 14 and 15. Figure 14 represents a typical continental station located in the polar continental air throughout the occurrence of

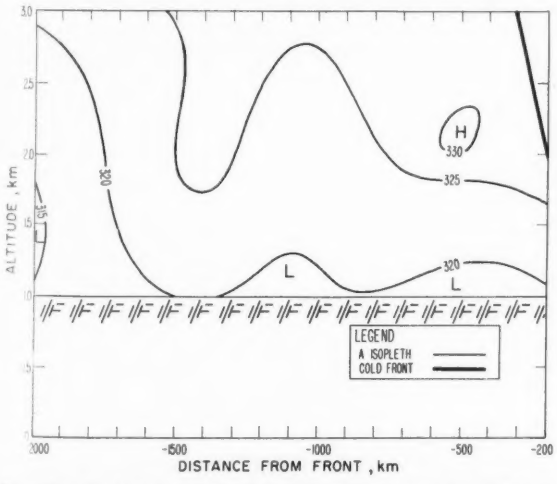


FIGURE 14. *The* A *unit pattern for Rapid City, S. Dak., expressed as a function of the distance from the front.*

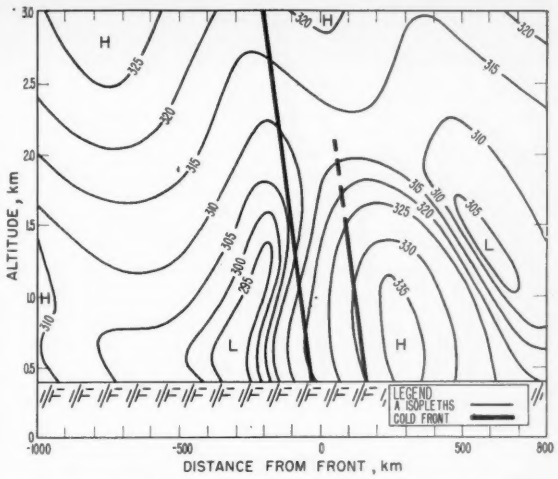


FIGURE 15. *The* A *unit pattern for Oklahoma City, Okla., expressed as a function of the distance from the front.*

the storm. The essential feature here is the absence of detail of A structure due to the presence of a uniform airmass over this station. Compare this with the epoch chart for Oklahoma City (fig. 15) where the structure of the idealized model is clearly reflected by the prefrontal A unit high, strong gradient across the frontal zone, and the A unit low behind the front. This rather fortuitous agreement is felt to be due to the strategic location of Oklahoma City with respect to the motion of the contrasting airmasses about the polar front. That is, this epoch chart represents a point of confluence of virtually unmodified polar continental and tropical maritime air.

5. Conclusions

The major conclusion of the present study is that an exponential correction to the refractive-index height distribution allows airmass properties to be more clearly seen than is allowed by previous techniques.

By the use of this exponential correction one may construct an idealized refractive-index field about a frontal transition zone that clearly shows the temperature and humidity contrasts of the different airmasses.

Further, when this technique is applied to the analysis of an actual storm, it does indeed highlight airmass differences.

6. References

[1] P. A. Sheppard, The structure and refractive index of the atmosphere, Rept. Conf. Meteorol. Factors Radio Wave Propagation (1946).

[2] N. G. Gerson, Variations in the index of refraction of the atmosphere, Geofisica Pura E Applicata **13,** 88 (1948).

[3] Andree Perlat, Meteorology and radioelectricity, L'Onde Elec. **28,** 44 (1948).

[4] D. R. Hay, Airmass refractivity in central Canada, Can. J. Phys. **43,** 12 (1958).

[5] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** 10 (1955).
[6] K. A. Norton, Point-to-point radio relaying via the scatter mode of tropospheric propagation, IRE Trans. **CS–4,** 1 (1956).
[7] R. E. Gray, The refractive index of the atmosphere as a factor in tropospheric propagation far beyond the horizon, IRE Natl. Conv. Record Pt. 1, 3 (1957).
[8] E. K. Smith and S. Weintraub, The constants in the equation for atmospheric refractive index at radio frequencies, Proc. IRE **41,** 1035 (1953).
[9] B. R. Bean and J. D. Horn, Radio refractive-index climate near the ground, J. Research NBS **63D,** (1959).
[10] W. A. Arvola, Refractive index profiles and associated synoptic patterns, Bull. Am. Meteorol. Soc. **38,** 212 (1957).
[11] B. R. Bean and L. P. Riggs, On the synoptic variation of the radio refractive index, J. Research NBS **63D,** (1959).
[12] C. L. Godske, T. Bergeron, and R. C. Bundgaard, Dynamic meteorology and weather forecasting, p. 526 (American Meteorological Society and Carnegie Institute of Washington, D.C., 1957).
[13] H. B. Byers, General meteorology, p. 327 (McGraw-Hill Book Co., Inc., New York, N.Y., 1944).
[14] F. A. Berry, E. Bollay, and Norman F. Beers, Handbook of meteorology, p. 638 (McGraw-Hill Book Co., Inc., New York, N.Y., 1945).
[15] J. B. Smyth and L. G. Trolese, Propagation of radio waves in the troposphere, Proc. IRE **35,** 1198 (1947).
[16] B. R. Bean and G. D. Thayer, On models of the atmospheric radio refractive index, Proc. IRE **47,** 740 (1959).

BOULDER, COLO. (Paper 63D2–21.)

Publications of the National Bureau of Standards

Selected Abstracts

Source function in a nonequilibrium atmosphere: III. The influence of a chromosphere, J. T. Jefferies and R. N. Thomas, *Astrophys. J.* **129**, *401* (*March 1959*).

Methods previously developed by the authors (Astrophys. J. **125**, 260 (1957) and **127**, 667 (1958)) are applied to investigate the depth dependence of the source function for resonance lines in an atmosphere having a chromospheric distribution of T_e superposed upon a photospheric one. The derived behavior of $S_L(\tau)$ for the neutral and ionized metals differs and mimics the observed behavior of such lines. The hydrogen Balmer lines should behave like the neutral metals, and the predicted behavior agrees with earlier empirical results obtained by the authors.

Transmission of power in radio propagation, J. R. Wait, *Electronic Radio Engr.* **36**, *146* (*April 1959*).

The theoretical foundations of the transmission-loss concept are discussed. It is emphasized that the influence of the ground on the input resistances of the transmitting and receiving aerials must be accounted for. The concept of "propagation loss" is introduced and its usefulness is pointed out.

Electromagnetic radiation from cylindrical structures, J. R. Wait (*Pergamon Press Inc., New York, N. Y. 1959*).

A comprehensive and self-contained account of the theory of electromagnetic sources on and near cylindrical surfaces is presented. While none of the forms considered duplicates the shape of actual bodies such as aircraft and missiles, the derived results do provide considerable insight into the behavior of flush-mounted and stub-type microwave antennas in actual systems. In each section, the salient features of the analytical results are presented in graphical form. An exhaustive and up-to-date review of the relevant literature in the field is given. This part of the work was sponsored by Commission VI of the International Scientific Radio Union. The contents include: Axial dipole within a corner reflector and slots on wedges, The fields of axial dipoles in the presence of a circular conducting cylinder, Slots on a cylindrically-tipped wedge, Transverse half-wave slots on isolated cylinders, Conductance of axial slots on infinite cylinders and wedges, Residue series representation for the fields of an axial slot, Fields of the axially slotted cylinder in terms of Fock functions, Residue series representation for circumferential slots on cylinders, Fields of current distributions in wedge regions, Radial electric dipoles in cylindrical wedge regions, Field of a slot on an elliptic cylinder, and Calculated radiation characteristics of slots in metal sheets.

Theoretical multiplet strengths, F. Rohrlich, *Astrophys. J.* **129**, *441* (*March 1959*).

General formulas are derived for all multiplet strengths of astrophysical interest, covering all transitions of the types $l^n.l'\text{-}l^n.l''$, $l^n\text{-}l^{n-1}.l'$, $l^n.l'\text{-}l^{n-1}.l'^2$, $l^n l'.l''\text{-}l^n.l'^2$, and $l^n l''.l'\text{-}l^n.l'^2$. The expressions are given in terms of the fractional parentage and Racah coefficients, which are both tabulated. With their aid, all transitions between multiply occurring terms can be separated according to parent and, if necessary, grandparent terms. No normalization factors are needed. The results greatly enlarge the scope defined by the available tables of multiplet strengths and also permit the correction of some errors in some of these tables.

Other NBS Publications*

Journal of Research, Section A. Physics and Chemistry, Volume 63A, No. 1, July–August 1959. 70 cents. Approximate issue date, August 3, 1959.

Description and analysis of the first spectrum of iodine. C. C. Kiess and C. H. Corliss.

CH in the solar spectrum. Charlotte E. Moore and H. P. Broida. Infrared studies in the 1, to 15-micron region to 30,000 atmospheres. C. E. Weir, E. R. Lippincott, A. Van Valkenburg, and E. N. Bunting.

Phosphinoborine compounds: mass spectra and pyrolysis. L. A. Wall, S. Straus, R. E. Florin, F. L. Mohler, and P. Bradt.

Experimental and theoretical study of kinetics of bulk crystalization in poly (chlorotrifloroethylene). J. D. Hoffman, J. J. Weeks, and W. M. Murphey.

Journal of Research, Section A. Physics and Chemistry, Volume 63A, No. 2, September–October 1959. 70 cents. Approximate issue date, October 1, 1959.

Reflection of fast neutrons from water. Martin J. Berger and John W. Cooper.

Some vibrational-rotational bands of deuterated methanes. Harry C. Allen, Jr., and Earle K. Plyler.

Precise coulometric titration of acids and bases. John K. Taylor and Stanley W. Smith.

Heat of formation of titanium tetraiodide. W. H. Johnson, A. A. Gilliland, and E. J. Prosen.

Reactions of pentafluorohalobenzenes. Walter J. Pummer and Leo A. Wall.

Tritium-labeled compounds I. Radioassay of tritium-labeled compounds in "Infinitely Thick" films with a windowless, gas-flow, proportional counter. Horace S. Isbell, Harriet L. Frush, and Ruth Peterson.

Tritium-labeled compounds II. General-purpose apparatus, and procedures for the preparation, analysis, and use of tritium oxide and tritium-labeled lithium borohydride. Horace S. Isbell and Joseph D. Moyer.

Preparation of trichloride and tetrachloride of molybdenum. Dwight E. Couch and Abner Brenner.

Infrared spectra of asphalts. Burton D. Beitchman.

Selected abstracts of publications by the National Bureau of Standards staff.

Journal of Research, Section B. Mathematics and Mathematical Physics, Volume 63B, No. 1, July–September 1959. 75 cents. Approximate issue date, September 8, 1959.

Relations between summation methods and integral transformations. W. Greub.

On a modification of Watson's lemma. F. Oberhettinger.

** Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (Order by PR number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the author.*

Principal submatrices of a full-rowed non-negative matrix. K. Goldberg.

Zeros of certain polynominals. A. J. Goldman.

Tables of transport integrals: A supplement. W. M. Rogers, W. J. Hall, and R. L. Powell.

Lens design: A new approach. O. N. Stavroudis.

Analytical integration of differential equation for water storage. V. M. Yevdjevich.

Compressible turbulent boundary layers with heat transfer and pressure gradient in flow direction. A. Walz.

Journal of Research, Section C. Engineering and Instrumentation, Volume 63C, No. 1, July–September 1959. 75 cents. Approximate issue date, August 17, 1959.

Creep of cold-drawn nickel. W. D. Jenkins and C. R. Johnson.

Friction and endurance of prelubricated and unlubricated ball bearings at high speeds and extreme temperatures. H. S. White.

Effect of light and water on the degradation of asphalt. L. R. Kleinschmidt and H. R. Snoke.

Current and potential relations for the cathodic protection of steel in a high resisitivity environment. W. J. Schwerdtfeger.

A tilting air-lubricated piston gage for pressures below one-half inch of mercury. U. O. Hutton.

Compact multianvil wedge type high-pressure apparatus. E. C. Lloyd, U. O. Hutton, and D. P. Johnson.

A coulometric-titration coulometer. S. W. Smith and J. K. Taylor.

Electron beam magnetometer. L. Marton, L. B. Leder, J. W. Coleman, and D. C. Schubert.

A refined X-band microwave microcalorimeter. G. F. Engen.

On the theory of fading properties of a fluctuating signal imposed on a constant signal, H. Bremmer, C599, 25 cents.

Stark broadening functions for the hydrogen lines, A. B. Underhill and J. H. Waddell, C603, 65 cents.

Maximum permissible body burdens and maximum permissible concentrations of radionuclides in air and in water for occupational exposure, (supersedes H52) H69, 35 cents.

Wind pressures in various areas of the United States, G. N. Brekke, BMS152, 15 cents.

Fractional factorial experiment designs for factors at three levels, W. S. Connor and M. Zelen, AMS54, 30 cents.

Tables of osculatory interpolation coefficients, H. E. Salzer, AMS56, 30 cents.

A double-pulse total-absorption fast neutron spectrometer, L. J. Nicastro and R. S. Caswell, TN1 (PB151360).

World maps of F2 critical frequencies and maximum usable frequency factors, D. H. Zacharisen, TN2 (PB151361), $3.50.

Detwinning quartz crystals, F. P. Phelps, TN3 (PB151362), 50 cents.

The vapor pressures of some hydrocarbons in the liquid and solid state at low temperatures, W. T. Ziegler, TN4 (PB151363), 75 cents.

Tables of median hourly values of the cosine of the solar zenith angle (χ) for 35 locations, R. F. McDuffie, TN5 (PB151364), $1.25.

An analysis of propagation measurements made at 418 Mc well beyond the radio horizon, H. B. Janes, J. C. Stroud, and M. T. Decker, TN6 (PB151365), $2.25.

Low- and very low-radiofrequency tables of ground wave parameters for the spherical earth theory: The roots of Riccati's differential equation (supplementary numerical data for NBS Circ. 573), J. R. Johler, L. C. Walters, and C. M. Lilley, TN7 (PB151366), $2.25.

Recently developed microwave impedance standards and methods of measurement, R. W. Beatty and D. M. Kerns, IRE Trans. Instrumentation 1–7, nos. 3–4 (1958).

The origin of the current nomenclature for the ionospheric layers, R. Silberstein, J. Atmospheric and Terrest. Phys. **13,** 382 (1959).

A survey of the present knowledge of sporadic-E ionization. J. A. Thomas and E. K. Smith. J. Atmospheric and Terrest. Phys. **13,** 295 (1959).

Downcoming radio waves; measurement of characteristics. Electronic and Radio Engr. **36,** 106 (1959).

Dense subgraphs and connectivity. R. E. Nettleton, K. Goldberg, and M. S. Green, Can. J. Math, **11,** 262 (1959).

Experiments with the cell Bi, Bi_2O_3, $ZnCl_2/ZnCl_2$/ $ZnCl_2$, Zn between 450° and 510°, R. E. Wood, J. Phys. Chem. **63,** 525 (1959).

Rheology of polyisobutylene. III. Elastic recovery, non-Newtonian flow, and molecular weight distribution, H. Leaderman, R. G. Smith, and L. C. Williams, J. Polymer Sci. **36,** 233 (1959).

The plasma jet as a spectroscopic source, M. Margoshes and B. F. Scribner, Spectrochim. Acta. 138 (1959).

Thermal degradation of polyethylene oxide and polypropylene oxide, S. L. Madorsky and S. Straus, J. Polymer Sci. **36,** 183 (1959).

U^{235} fission product decay spectra at various times after fission, A. T. Nelms and J. W. Cooper, Health Phys. **1,** 427 (1959).

Contribution to the theory of Brownian motion, R. W. Zwanzig, Physics of Fluids **2,** 1, 12 (Jan-Feb. 1959).

Low-temperature X-ray studies of Rice's blue material, L. H. Bolz, F. A. Mauer, and H. S. Peiser, J. Chem. Phys. **30,** 1, 349 (1959).

Contribution of annihilation radiation to the gamma-ray flux in lead, M. J. Berger, J. H. Hubbell, and I. H. Reingold, Phys. Rev. **113,** 3, 857 (1959).

On the electromagnetic response of an imperfectly conducting thin dyke, J. R. Wait, Geophysics **24,** 1, 167 (1959).

Temperature dependence of the breakdown field of barium titanate, P. H. Fang and W. S. Brower, Phys. Rev. **113,** 2, 456 (1959).

A re-examination of the crystal structures of α and β nitrogen, L. H. Bolz, M. E. Boyd, F. A. Mauer, and H. S. Peiser, Acta Cryst. **12,** Part 3, 247 (1959).

A study of the morphology of ionospheric storms, S. Matsushita, J. Geophys. Research **64,** 3, 305 (1959).

Energetic species trapped at 4.2° K from gaseous discharges, G. J. Minkoff, F. I. Scherber, and J. S. Gallagher, J. Chem. Phys. **30,** 3, 753 (1959).

Oscillatory currents in the ionosphere (abstract), Astron. J. **64,** 56 (1959).

A new method of checking the consistency of precedence matrices, R. B. Marimont, J. Assoc. Computing Machinery **6,** 2, 164 (1959).

Comments on the structure of 1,1,1-trifluoro-2-butyne, V. W. Laurie, J. Chem. Phys. **30,** 4, 1101 (1959).

Effect of heat-treatment on the constitution and mechanical properties of some hydrated aluminous cements, J. Am. Ceram. Soc. **42,** 4, 185 (1959).

Electronic quenching of OH ($^2\Sigma^+$) in flames and its significance in the interpretation of rotational relaxation, T. Carrington, J. Chem. Phys. **30,** 4, 1087 (1959).

IGY observations of F-layer scatter in Far East, R. Bateman, J. W. Finney, E. K. Smith, L. H. Tveten, and J. M. Watts, J. Geophys. Research **64,** 4, 403 (1959).

Reflectors for a microwave Fabry-Perot interferometer, W. Culshaw, IRE Trans. Microwave Theory Tech. **MTT-7,** 221 (1959).

Testing and operation of ball bearings submerged in liquefied gases, K. B. Martin and R. B. Jacobs, ASLE Trans. **2,** 1, 101 (1959).

Diurnal change of ionospheric heights deduced from phase velocity measurements at VLF (letter), J. R. Wait, Proc. IRE **47,** 998 (1959).

Green coronal line intensity and geomagnetism, C. Warwick, J. Geophys. Research **64,** 5, 527 (1959).

Models of the atmospheric radio refractive index, B. R. Bean and G. D. Thayer, Proc. IRE **47,** 740 (1959).

Sporadic *E* at VHF in the USA, R. M. Davis, Jr., E. K. Smith, and C. D. Ellyett, Proc. IRE **47,** 5, 762 (1959).

Recent developments in the field of microwave power measurements at the National Bureau of Standards, G. F. Engen, IRE Trans. Instrumentation **1-7,** Nos. 3-4 (1958).

Experiments on the β decay of polarized nuclei, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Physica. **24,** S64 (1958), No. 40 (1959).

Standards and measurements for electronics, A. V. Astin, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 134 (1958).

A free precession determination of the proton gyromagnetic ratio, P. L. Bender and R. L. Driscoll, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 176 (1958).

Atomic beam sources and the standard of length, K. G. Kessler, R. L. Barger, W. G. Schweitzer, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 181 (1958).

The ammonia maser as an atomic frequency and time standard, R. C. Mockler, J. Barnes, R. Beehler, H. Salazar, and L. Fey, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 201 (1958).

AC-DC transfer instruments for current and voltage measurements, F. L. Hermach, IRE Trans. Instrumentation **1-7,** No. 3-4, 235 (1958).

New apparatus at the National Bureau of Standards for absolute capacitance measurement, M. C. McGregor, J. F. Hersh, R. D. Cutkosky, F. K. Harris, and F. R. Kotter, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 253 (1958).

High-frequency standards of the electronic calibration center, NBSBL, M. C. Selby, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 262 (1958).

High-frequency impedance standards at the National Bureau of Standards, R. C. Powell, R. M. Jickling, and A. E. Hess, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 270 (1958).

A dry, static calorimeter for RF power measurement, P. A. Hudson and C. M. Allred, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 292 (1958).

Classification and nomenclature for standards of measurement, A. G. McNish, IRE Trans. Instrumentation **1-7,** Nos. 3-4, 371 (1958).

A note on the computation of x^2, M. G. Natrella, Am. Statistician **13,** No. 1, 20 (1959).

On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. II, A. M. Ostrowski, Arch. Rational Mech. Anal. **2,** No. 5, 423 (1959).

Un nouveau critere d'univalence des transformations dans un R^n. A. M. Ostrowski, Compt. rend. acad. sci. **248,** 348 (1959).

Modulation studies for VHF ionospheric scattering, J. W. Koch, Record of National Symposium on Extended Range and Space Communications, 114 (Washington, D.C., Oct. 6-7, 1958).

1958 Critique of VHF ionospheric scatter communication (Survey Paper), R. C. Kirby, Record of National Sumposium on Extended Range and Space Communications, 90 (Washington, D.C., October 6-7, 1958).

The diode reactance modulator, G. F. Montgomery, Commun. Electronics, No. 40, 1 (Jan. 1959).

A statistician and the post office: A case history in operations research, N. C. Severo, Trans. 1959 Conf. Admin. Appl. Div. of ASQC (Milwaukee, Wis., 1959).

Beta-gamma correlations from polarized nuclei, Chapter III (Interactions Faibles), E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Compt. rend. Congr. Intern. Phys. Nucleaire, Paris, July 1958 (1959).

The effect of echo on the operation of high-frequency communication circuits, D. K. Bailey, IRE Trans. Antennas and Propagation **6,** No. 4, 325 (1958).

An improved 8-hydroxyquinoline method for the determination of magnesium oxide in portland cement, H. A. Berman, ASTM Bull. No. 237, 51 (1959).

Some radiological applications of gamma-ray transport theory, M. J. Berger and L. V. Spencer, Radiation Research **10,** No. 5, 552 (1959).

Degradation of cellulose acetate films, W. K. Wilson and B. W. Forshee, SPE Journal **15,** No. 2 (1959).

Multi-factor experiments, M. Zelen and W. S. Connor, Ind. Quality Control, **15,** No. 9, 1 (1959).

Science Fairs, A. T. McPherson, ASTM Bull. No. 237, **29** (1959).

Selection of glasses for three-color achromats, R. E. Stephens, J. Opt. Soc. Am. **49,** No. 4, 398 (1959).

Note on measurement of sine-wave response of lenses, R. E. Stephens, J. Opt. Soc. Am. **49,** No. 4, 413 (1959).

The elasticity of rubber, L. A. Wood, Rubber Chem. and Technol, **31,** No. 5, 959 (1958).

Thermal decomposition of poly(vinyl chloride), R. R. Stromberg, S. Strauss, and B. G. Achhammer, J. Polymer Sci. **35,** No. 129, 355 (1959).

Thermal emf of some thermometric alloys, M. D. Bunch, R. L. Powell, and R. J. Corruccini, In Low-Temperature Physics and Chemistry; Proc. 5th Internatl. Conf., Univ. of Wis., Madison, Wis., 1957, 484 (1958).

Thermal and electrical conductivity of aluminum and aluminum alloys, W. J. Hall, H. M. Roder, and R. L. Powell, In Low-Temperature Physics and Chemistry; Proc. 5th Internatl. Conf., Univ of Wis., Madison, Wis., 1957, 389 (1958).

Thermal and electrical conductivity of pure copper, H. M. Roder, R. L. Powell, and W. J. Hall, In Low-Temperature Physics and Chemistry; Proc. 5th Internatl. Conf. Univ. of Wis., Madison, Wis., 1957, 364 (1958).

Formation of ozone from atomic oxygen at low temperatures, R. A. Ruehrwein and J. S. Hashman, J. Chem. Phys. **30,** No. 3, 823 (1959).

The electromotive force method of determining ionization constants: Ionization of bisulfate ion, W. J. Hamer, (Book) The Structure of Electropytic Solutions, edited by W. J. Hamer, 236 (John Wiley & Sons, Inc., New York, N.Y., 1959).

Kinetics of D atom reactions with H_2, CH_4, and D_2CO, F. Klein, J. R. McNesby, M. D. Sheer, and L. J. Schoen, J. Chem. Phys. **30,** No. 1, 58 (1959).

Natural and synthetic rubbers, M. Tryon and F. J. Linnig, Anal. Chem. Appl. Revs. **31,** Part II, No. 4, 767 (1959).

Molecular complexes and their spectra. IX, Infrared absorption by iodine in its pyridine complexes and in benzene, E. K. Plyler and R. S. Mulliken, J. Am. Chem. Soc. **81,** 828 (1959).

The emission of HeI, λ 10830, during the great flare of August 26, 1958, E. Tandberg-Hanssen, W. Curtis, and K. Watson, Astrophys. J. **129,** No. 1, 238 (1959).

Direct-coupled power amplifier for cryostat heating control, R. D. Goodwin and J. R. Purcell, Rev. Sci. Instr. **28,** No. 7, 581 (1957).

Tacan coverage and channel requirements, M. T. Decker, IRE Trans. Aeronaut. Navigational Electronics **ANE–4,** 3, 135 (1957).

A phenomenological theory of induced electrical polarization, J. R. Wait, Can. J. Phys. **36,** 1634 (1958).

A study of VLF field strength data: Both old and new, J. R. Wait, Geofis, pure e Appl. **41,** 73 (1958/III).

Near infrared solar radiation measurements by balloon to an altitude of 1,000,000 feet, D. M. Gates, D. G. Murcray, C. C. Shaw, and R. J. Herbold, J. Opt. Soc. Am. **48,** 12, 1010 (1958).

Products of cool flame oxidation of propane-2-^{13}C, R. E. Ferguson and C. R. Yokley, Seventh Symposium (International) on Combustion (London and Oxford, 28 Aug.–3 Sept. 1958) Butterworths Scientific Publications, London (1958).

Emission spectra from mixtures of atomic nitrogen and organic substances, N. H. Kiess and H. P. Broida, Seventh Symposium (International) on Combustion (London and Oxford, 28 Aug.–3 Sept. 1958) Butterworths Scientific Publications, London (1958).

United States planning for the large scale liquefaction and transportation of helium, D. B. Mann and B. W. Birmingham, Physica **24,** S144 (1958)

Transport properties of dilute alloys of copper, R. L. Powell, W. J. Hall, and H. M. Roder, Physica **24,** S176 (1958).

The application of the theory of stochastic processes to chemical kinetics, E. W. Montroll and K. E. Schuler, (Book) Advances in Chemical Physics, Vol. 1 (Interscience Publishers, Inc., New York, N.Y., 1958).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Radio-Refractive-Index Climate Near the Ground

B. R. Bean and J. D. Horn

(February 6, 1959)

The radio refractive index of air is a function of atmospheric pressure, temperature, and humidity and is found to vary in a systematic fashion with climate. It was found that the surface value of the refractive index may be estimated four to five times more accurately from charts of reduced-to-sea-level values than from similar sized charts of surface index. Worldwide maps of 5-year means of this reduced value are presented for the months of February and August, for the minimum monthly mean value of the year and for the range of monthly mean values. Year-to-year variation of monthly means is also considered. Applications of these data to the prediction of radio field strengths indicate a possible 30-decibel difference in median level of identically equipped tropospheric communications systems due to climate alone.

1. Introduction

The radio refractive index of air, n, is a function of atmospheric pressure, temperature and humidity, thus combining in one parameter, three of the normal meteorological elements used to specify climate. In the following sections we will examine the variability of n during different seasons of the year and in differing climatic regions. The systematic dependence of n upon station elevation will make it necessary to consider a method of expressing n in terms of an equivalent sea-level value in order to see more clearly the actual climatic differences of the various parts of the world. After a consideration of the n climate of the world, the application of this information to such practical problems as the prediction of radio field strength and the refraction of radio waves will be discussed.

2. Presentation of Basic Data

Near the surface of the earth, for vhf and uhf frequencies, n is a number of the order of 1.0003. Since, for air, n never exceeds unity by more than a few parts in 10^{-4}, it is convenient to consider the climatic variation of n in terms of

$$N=(n-1)10^6=\frac{77.6}{T}\left(P+\frac{4810e_s RH}{T}\right), \qquad (1)$$

where P is the total atmospheric station pressure in millibars (mb), RH is the relative humidity in percent, and e_s the saturation vapor pressure in mb at the temperature, T, in deg Kelvin. The values of the constants in (1) were determined [1] [1] from a consideration of recent microwave and optical determinations of the refractive index of air and are considered to be accurate to 0.5 percent in N for frequencies up to 30,000 Mc in the ranges of temperature, pressure and humidity normally encountered. The notation N_s is used to indicate that (1) has been evaluated from standard surface weather observations.

To obtain long-term average values of N one should properly average individual observations over many years. This is difficult to do since, in general, only summaries of weather observations are readily available. However, long-term average values of temperature, pressure and humidity are available and may be converted into an "average" value of N. This "average" N differs from the true average since the intercorrelation of pressure, temperature, and humidity is neglected. This difference was examined by an analysis of 2 yr of weather records of the months of February and August at an arctic location (Fairbanks, Alaska), a temperate zone location (Washington, D.C.), and a tropical location (Swan Island, W.I.). These data, given in table 1, indicate that the difference between the two methods was never more than 1.5 N units and the average difference was less than 1 N unit which is small compared to commonly observed seasonal and geographic variations of 20 to 100 N units.

TABLE 1. *Two-year average value of* N_s *versus the value of* N_s *calculated from average temperature, pressure, and humidity*

	$\overline{N_s}$	$\overline{N_s}(\overline{P},\overline{T},\overline{RH})$	$\overline{N_s}-\overline{N_s}(\overline{P},\overline{T},\overline{RH})$
Fairbanks:			
February	314.0	313.0	1.0
August	320.5	320.0	0.5
Washington:			
February	305.5	304.5	1.0
August	356.0	354.5	1.5
Swan Island:			
February	362.0	362.5	0.5
August	387.5	388.0	.5
Average			0.83

[1] Figures in brackets indicate the literature references at the end of this paper.

On this basis it was decided to use the long term means given in the United Nations' monthly publication, *Climatic Data for the World*. This publication is particularly advantageous for our present study since it reports the fictitious value of the relative humidity needed to obtain the actual average vapor pressure from the saturated vapor pressure of the reported mean temperature [2].

Data from 306 weather stations were obtained in order to give reasonable geographical coverage. In general, 5 yr of record were obtained for each station from the period 1949 to 1958, preference being given to the years 1954 through 1958. A noticeable exception, however, was Russia for which only 1 yr of data (IGY) is reported in *Climatic Data for the World* and thus all charts are drawn with dashed contours for Russia. There are vast expanses of ocean for which there are no meteorological observing stations. Climatic atlases were utilized in order to present estimates of world climate in these locales. A reasonable coverage of the sparse data areas of the world was made by estimating temperature from sea surface isotherms [3] and humidity from charts of seasonal average depression of the wet bulb temperature [4]. Pressure was estimated for these locations from average winter and summer pressure charts.

When these data were converted to N [5] and charts prepared, a pronounced altitude dependence could be seen as in figure 1. Figure 1 plus the following charts of N variations across the United States are from an extensive N climatology now being prepared at the Central Radio Propagation Laboratory. Although the present study is primarily aimed toward worldwide variations, it is felt that the U.S. data better illustrate the height dependence of N_s and the subsequent reduction process employed. It is noted that the coastal areas display high values of N_s while the inland areas have lower values. There are low values of N_s corresponding to the Appalachian and Adirondack Mountains, a decrease with increasing elevation of the Great Plains until the lowest values are observed in the Rocky Mountain region and the high plateau area of Nevada. A corresponding gradient is observed from the West Coast eastward. Crosshatching encloses areas where the terrain changes so rapidly that it was felt the data were inadequate to obtain realistic contours of N_s.

The altitude dependence of N can be studied in terms of the "dry" and "wet" components of N. The dry term, D,

$$D=\frac{77.6P}{T}, \tag{2}$$

is proportional to air density and normally constitutes at least 60 percent of N.

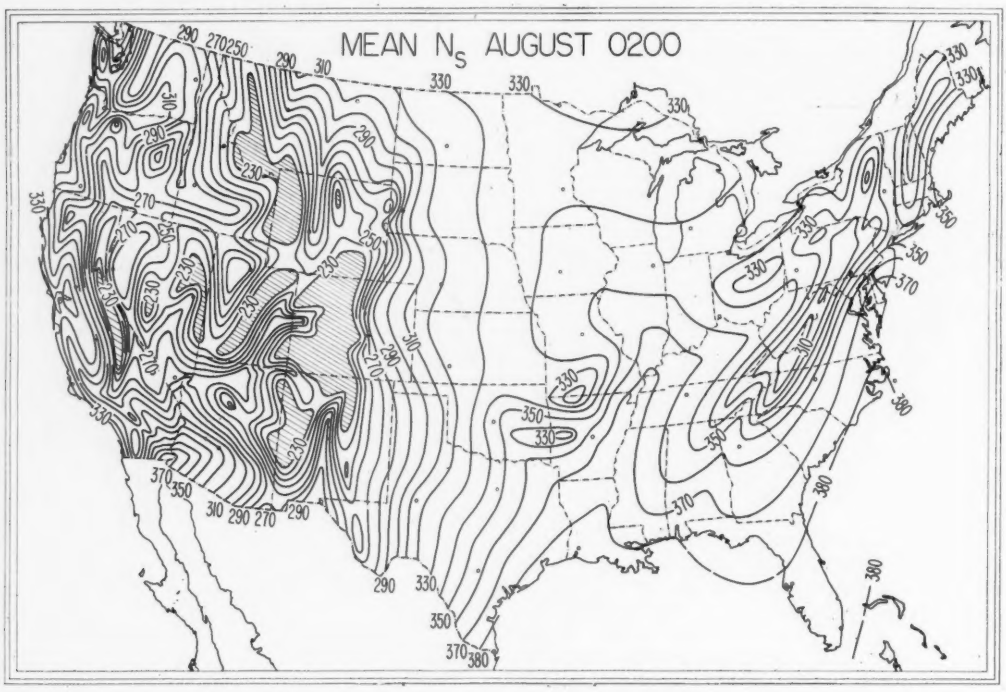


FIGURE 1. *Mean N_s, August 0200 local time.*

One would expect, from the hydrostatic equation in [6], that D would be an exponential function of height. Examination of D at the earth's surface versus station elevation shows this to be in fact true. By assuming an exponential decay between the value of N at sea level and 8 km the height coefficient of -0.1057/km was determined from the NACA dry standard atmosphere [7] and was adopted for general use. Thus, the station value, D_s, may be reduced to a sea-level value, D_0, by the relationship

$$D_0 = D_s \exp\{0.1057\, h\}, \tag{3}$$

where h is in kilometers. Values of D_0 obtained in this fashion as shown on figure 2 present a gradient that is remarkably free of detail compared to the N_s chart and is easily drawn for all areas of the country.

A similar investigation was made of the height dependence of the surface "wet" term, W_s, evaluated from

$$W = \frac{3.73 \times 10^5 e_s RH}{T^2}. \tag{4}$$

All the cases examined displayed low correlations of log W_s and height, indicating that W_s is not a marked exponential function of height. Thus W_s is plotted alone as on figure 3. Again, contours of W_s may be obtained easily for all sections of the country.

At this point in the reduction process there are two maps, one of D_0 and one of W_s. A further simplification is accomplished by introducing the approximation of reducing N_s by the dry term height correction to obtain a single reduced value, N_0;

$$N_0 = (D_s + W_s) \exp\{0.1057\, h\}. \tag{5}$$

Figure 4 gives the N_0 contours for the same time as the previous maps of D_0 and W_s. The N_0 maps are no more difficult to prepare than the W_s maps and have effectively removed the station height dependence of N_s. One might wonder at the advisability of arbitrarily reducing the wet term by the dry term correction. For the coastal areas of the country, where the exponential height correction factor is nearly one, this amounts simply to adding the D_0 and W_s maps while for the mountain areas, where the height correction factor is large, the W_s values are small with the result that the gradient of the N isopleths obtained from the D_0 and W_s maps is essentially maintained on the N_0 maps. As an example, for the series of maps under discussion, the $(D_0 + W_s)$ difference between Reno, Nev. (1,340-m elevation) and Oakland, Calif. (5.5-m elevation) is 21 N units, while the N_0 difference is 19 N units.

The effects of this correction on the worldwide values can be seen from figure 5 where N_s is plotted versus station elevation in km. A sample line illus-

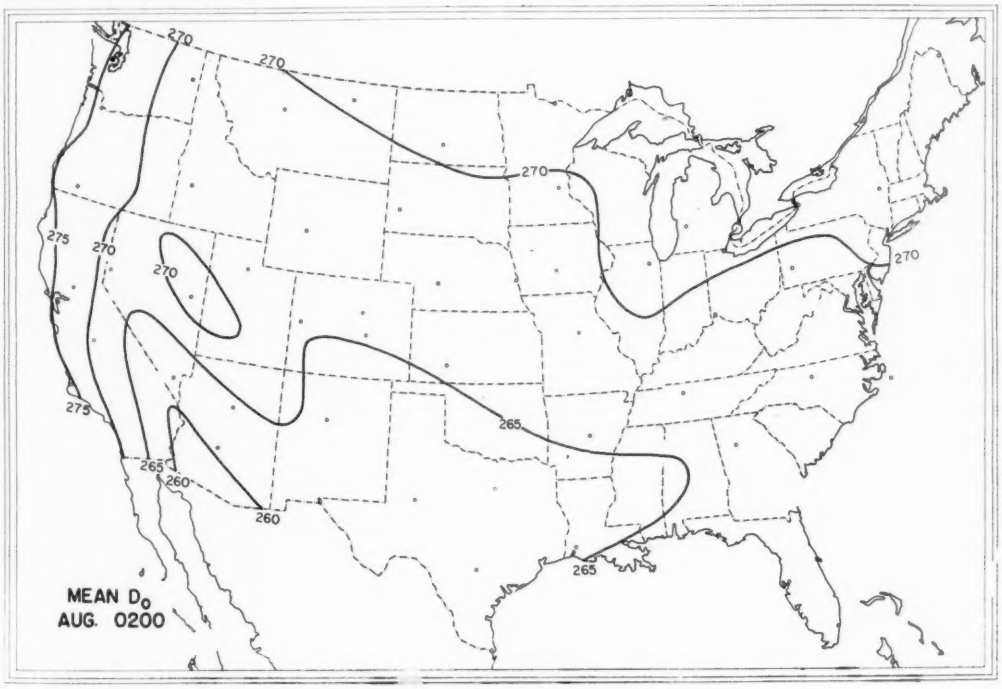


FIGURE 2. *Mean D_0, August 0200 local time.*

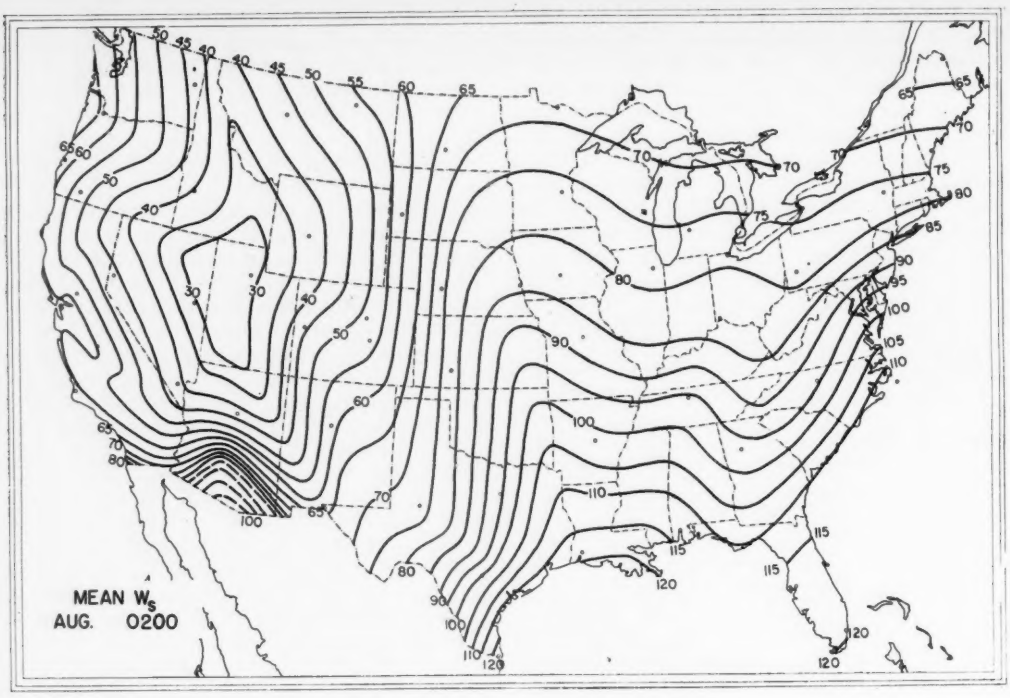


FIGURE 3. *Mean W_s, August 0200 local time.*

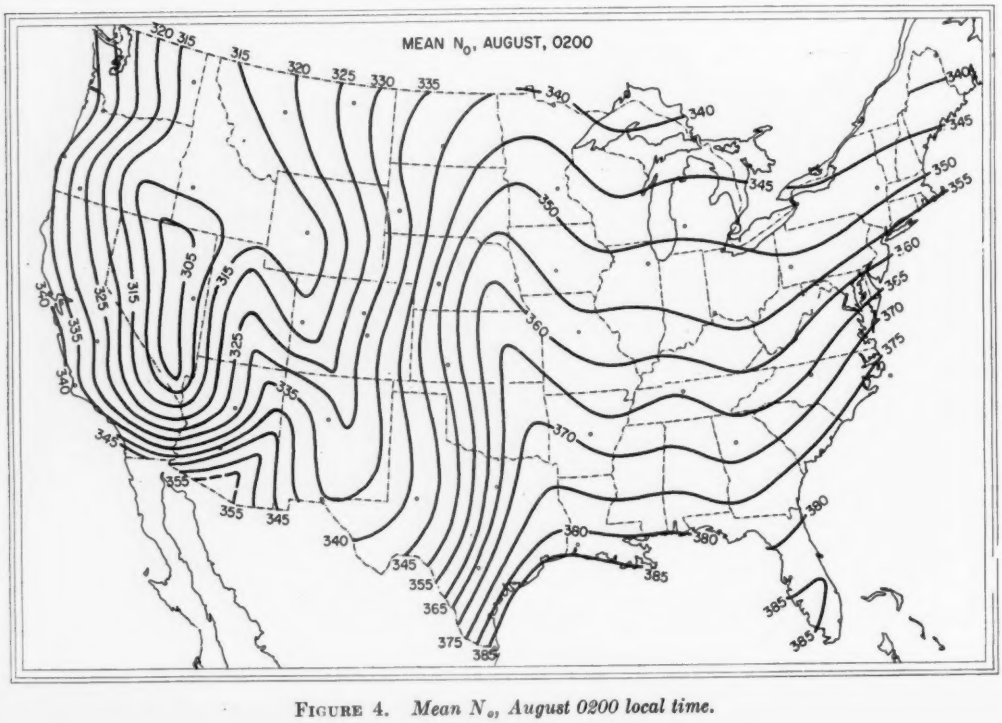


FIGURE 4. *Mean N_o, August 0200 local time.*

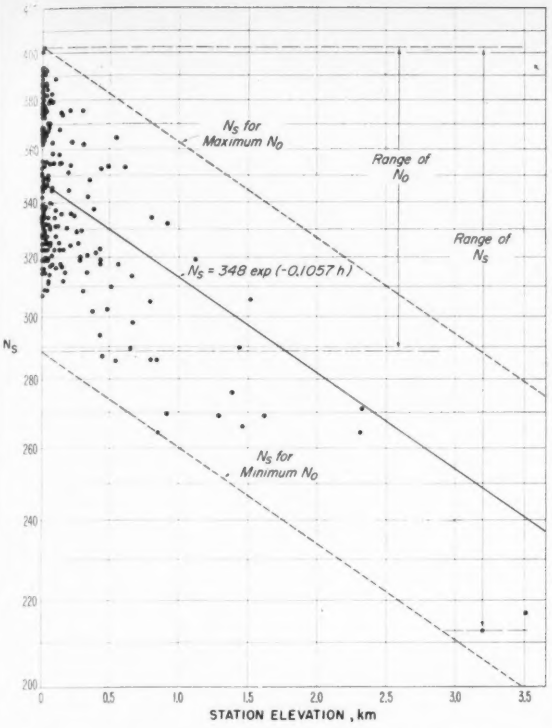


FIGURE 5. *Worldwide values of N_s versus height for August.*

trates the decay of N_s with height for $N_0=348$. The value of N_s for any other value of N_0 would be obtained from a line parallel to the $N_0=348$ line but having a zero intercept equal to the new value of N_0. The advantage of adopting N_0 is illustrated by the reduction in range from 190 N units for N_s to 115 N units for N_0, thus diminishing the number of contours of the resulting maps.

It would appear that by removing the influence of station elevation it would be more efficient to estimate N_s from N_0 charts rather than from charts of N_s. As a test of this hypothesis N_s and N_0 contour charts were prepared for both summer and winter from only 42 of the 62 U.S. Weather Bureau stations for which 8-yr means of N_s are available. The remaining 20 stations, distributed at random about the country, were used as a test sample by estimating their 8-yr mean value of N_s from the N_0 and N_s contours. Summertime examples of these charts are given by figures 6 and 7. Note that due to the reduced range of N, the N_0 charts are drawn every 5 N units as compared to the 10-N-unit contours of the N_s charts. The individual deviations of the values obtained from the contour maps with the actual 8-yr means are listed in table 2. By comparing the root mean square (rms) deviations of 10.7 N units in winter and 13.0 N units in summer obtained by estimating N_s from the N_s contours with the 2.7-N-unit rms of estimating N_s from N_0 contours, one concludes that it is at least 4 times more accurate to estimate N_s from N_0 contours than from those of N_s. An inspection of the individual deviations in table 2 indicates that the N_0 contour method is particularly efficient at elevations in excess of 1,200 m or where the terrain is changing rapidly with respect to horizontal distance. As a further practical consequence one notes the remarkable similarity between the N_0 contours of figures 4 and 7, even though the latter contours were derived from only two-thirds of the original data. This indicates that any desired level of accuracy may be maintained with fewer stations (and less expense) by the use of N_0.

TABLE 2.—*Deviations of estimated 8-yr means of N_s (calculated from contour charts) from actual 8-yr means for 20 randomly distributed U.S. Weather Bureau stations*

Test station	Height	February 1400			August 0200		
		Actual 8-yr mean value of N_s	Deviation [a]		Actual 8-yr mean value of N_s	Deviation [a]	
			N_s map	N_0 map		N_s map	N_0 map
	Meters	*N units*	*N units*	*N units*	*N units*	*N units*	*N units*
Sacramento, Calif.	7	315.6	7.6	0.8	329.6	1.0	−1.8
Portland, Oreg.	8	316.2	4.2	−.5	337.7	19.7	6.0
San Diego, Calif.	11	314.2	−5.8	−2.4	348.1	16.1	3.5
Mobile, Ala.	66	326.6	6.6	4.8	376.0	6.0	0.6
Fresno, Calif.	86	310.6	9.6	3.4	326.2	5.2	4.2
Boston, Mass.	89	308.6	−6.4	−0.5	347.5	−7.5	−0.4
Grand Rapids, Mich.	210	304.4	0.4	−5.9	340.5	−1.5	.1
Columbia, Mo.	239	300.8	−.2	2.4	348.7	−2.3	−2.5
Minneapolis, Minn.	255	301.1	.1	0.6	338.5	−0.5	2.7
Cincinnati, Ohio	271	302.5	−.5	.7	344.1	−2.9	−2.8
Des Moines, Iowa	294	300.9	3.9	2.3	343.1	−1.9	−0.1
Pendleton, Oreg.	455	295.9	1.9	0.4	300.9	2.9	−3.1
Billings, Mont.	1,088	269.3	−2.3	.1	285.6	5.6	1.2
Burns, Oreg.	1,262	268.1	−23.9	−3.2	271.3	−15.7	−4.4
Salt Lake City, Utah	1,288	266.3	1.3	−0.8	279.5	8.5	4.4
Reno, Nev.	1,340	259.6	−20.4	−6.8	277.6	−29.4	1.6
Pocatello, Idaho	1,355	264.7	−2.3	0.4	269.7	−3.3	0.0
Denver, Colo.	1,625	244.9	−8.1	.7	276.6	−1.4	.3
Colorado Springs, Colo.	1,882	237.1	−15.9	−.6	272.4	−6.6	1.1
Flagstaff, Ariz.	2,131	237.8	−26.2	2.3	261.4	−36.6	−2.2
Root mean square deviation	.....	.....	10.7	2.7	.....	13.0	2.7

[a] Deviation equals the actual long-term mean minus the value obtained from map contours.

3. Worldwide Values of N_0

Mean values of N_0 were calculated at each of the 306 selected stations and charts were prepared for each month of the year. The charts for February and August are given on figures 8 and 9. It is seen that the values of N_0 for sea-level stations vary from 390 in the maritime tropical areas to 290 in the deserts and high plateaus. The interior of continents and mountain chains in middle latitudes are reflected by low values as compared to those of coastal areas. Further, such pronounced climatic details as the Indian monsoon and the effects of coastal mountain ranges blocking prevailing winds and producing rain shadows are indicated by these N_0 contours.

The annual variation of N_s is indicated on figure 10 by contours of the difference between the maximum and minimum monthly means observed throughout the year. It is quite remarkable how clearly climatic differences are evidenced by the

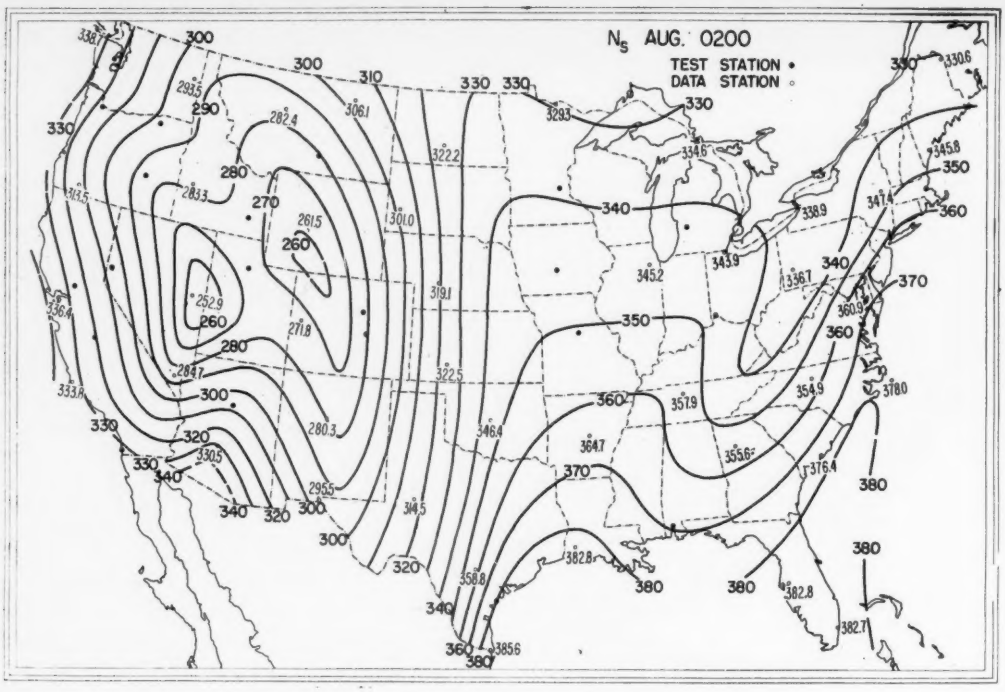


FIGURE 6. *Test chart of mean N_s August 0200 local time.*

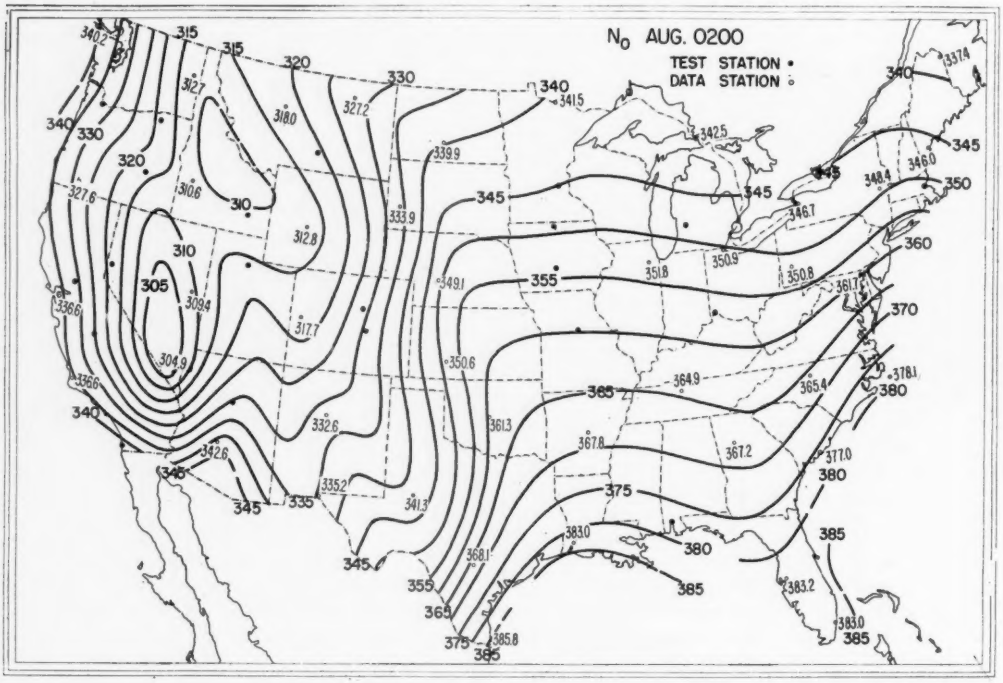


FIGURE 7. *Test chart of mean N_0 August 0200 local time.*

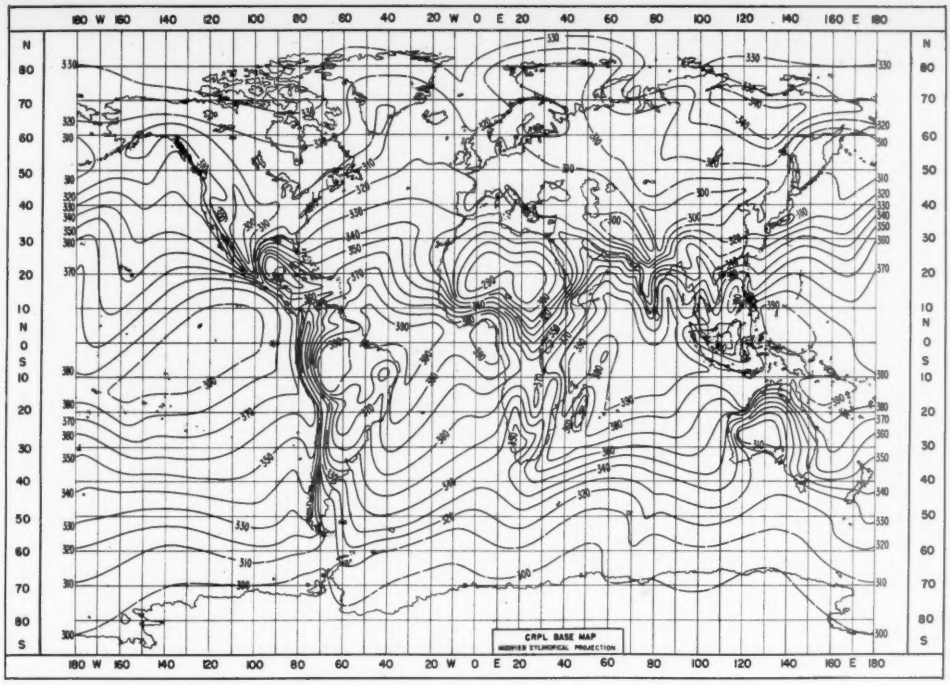


FIGURE 8. *Worldwide values of mean N_o for February.*

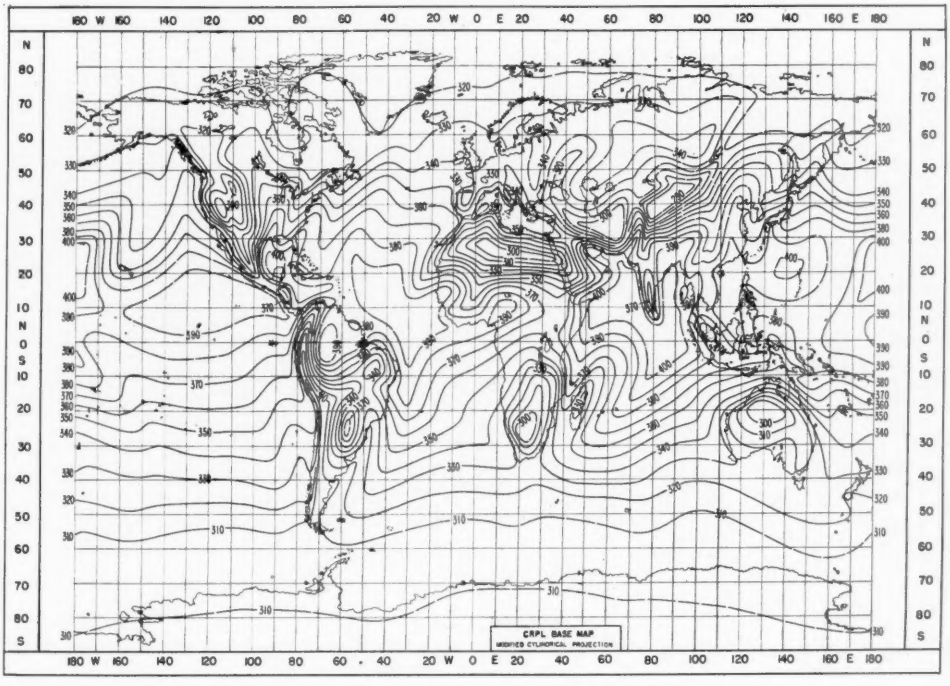


FIGURE 9. *Worldwide values of mean N_o for August.*

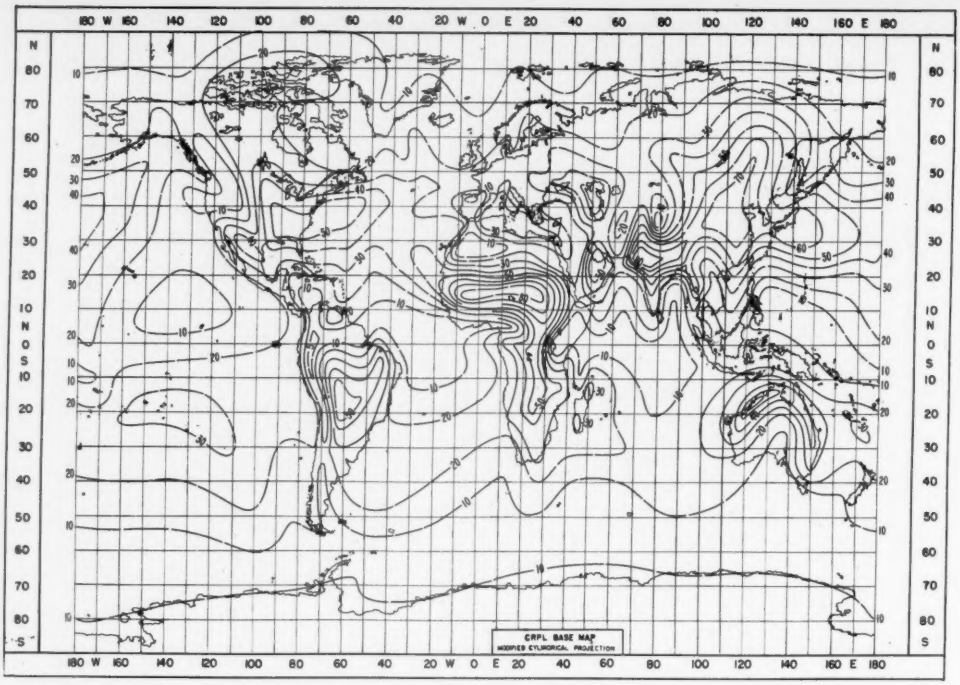


FIGURE 10. *Annual range of monthly mean N_s.*

yearly range of N_s. The prevailing transport of moist maritime air inland over the west coasts of North America and Europe is indicated by relatively small annual ranges (20 to 30 N units), while, for example, the east coast of the United States with a range of 40 to 50 N units or more reflects the invasion of that area from time to time by such diverse air-masses as arctic continental and tropical maritime. The largest annual ranges of N_s (90 N units) are observed in the Sudan of Africa and in connection with the Indian monsoon.

An additional N_0 map (fig. 11) was prepared from the minimum monthly mean value of N_s observed throughout the year to supplement the range map in order that an estimation might be made of both the minimum and maximum monthly mean N_s expected during the year.

A measure of the variability of the February and August mean values of N_s is given by monthly range maps (fig. 12 and 13) determined from monthly averages from 5 yr of data. Ranges are given by the maximum difference of the five individual monthly mean values. In contouring the two variability maps only those terrestrial regions having reasonable data coverage are included. Dashed contours are shown for areas of sparse or unreliable data. The general picture of the worldwide distribution of N_s variability is that of a number of continentally located cells of moderate range accompanied by somewhat random small-scale variations over ocean areas. Regions of large range, from 40 to as much as 70 N units, are present, however, in Australia and on islands of the adjoining oceans, on the African equatorial plateau near the Cameroons and in the Great Basin of the southwestern United States. Common to all these areas of large year-to-year variability, at least during the summer season, are high mean temperatures ranging from about 25° to 30° C, the variability being due to relatively small variations of humidity. It is felt that when a more dense network of stations is available for a longer period of record, say 10 yr, areas of high monthly variability are likely to be more extensive in tropical and desert areas than indicated on our present maps.

4. Climatic Classification by N_s

The annual cycle of N_s at each station was examined for the purpose of deriving similarities of climatic pattern. As one form of climatic classification, the annual mean value of N_s at each station was plotted versus the annual range at the station. When this was done, several distinct groupings of data seemed evident. These groupings, described in table 3, are intended to give a general idea of the geographic and climatic character of the majority of the stations found within given values of range and yearly mean of N_s.

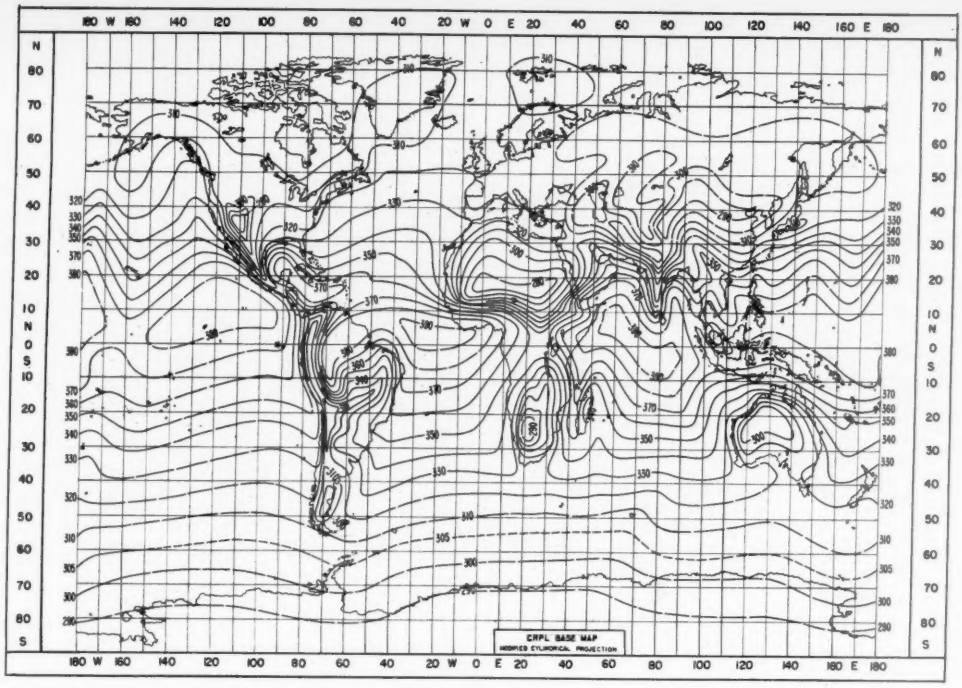


FIGURE 11. *Minimum monthly mean value of N_o.*

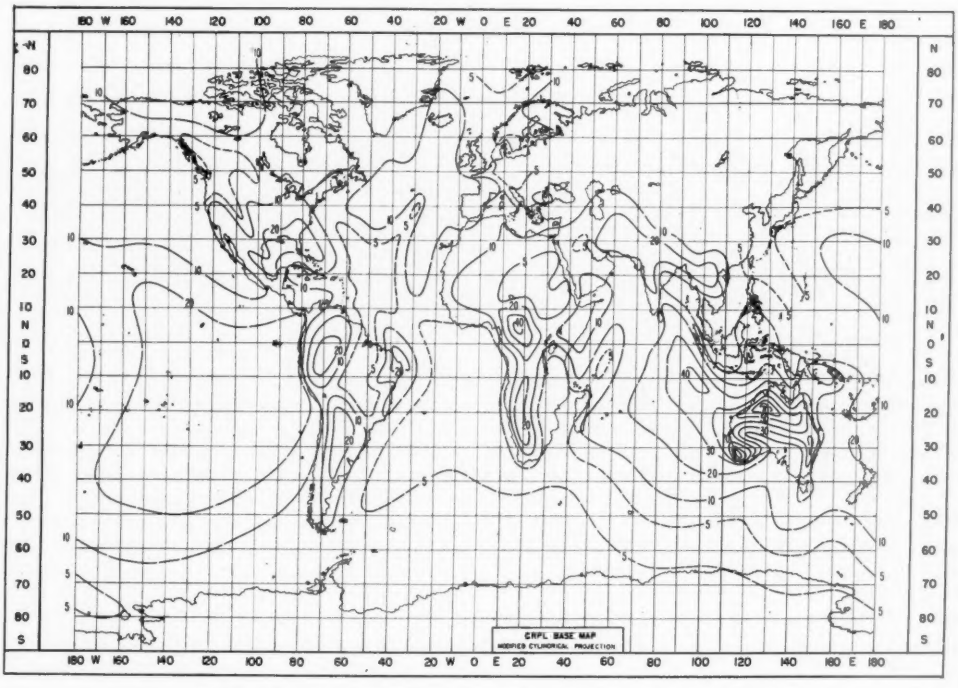


FIGURE 12. *Year-to-year range of monthly mean N_s for February.*

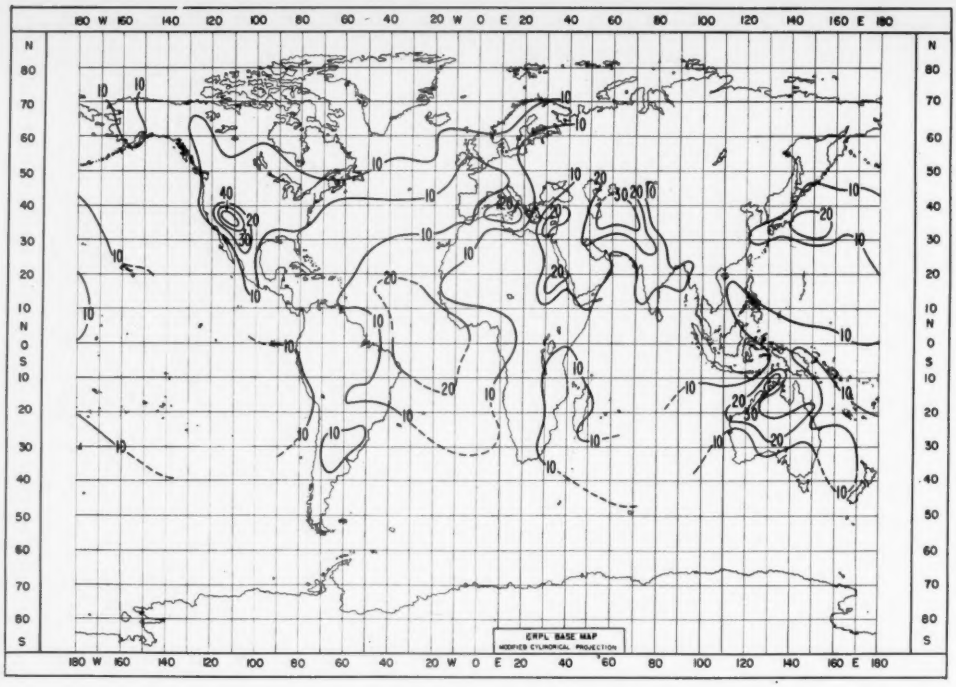


FIGURE 13. *Year-to-year range of monthly mean N_s for August.*

TABLE 3. *Characteristics of climatic types*

Type	Location	Annual mean N_s	Annual range of N_s	Characteristics
		N units	*N units*	
I Midlatitude-coastal	Near the sea or in lowlands on lakes and rivers, in latitude belts between 20° and 50°.	300 to 350	30 to 60	Generally subtropical with marine or modified marine climate.
II Subtropical-Savanna	Lowland stations between 30° N and 25° S lat., rarely far from the ocean.	350 to 400	30 to 60	Definite rainy and dry seasons, typical of Savanna climate.
III Monsoon-Sudan	Monsoon—generally between 20° and 40° N lat., Sudan—across central Africa from 10° to 20° N lat.	280 to 400	60 to 100	Seasonal extremes of rainfall and temperature.
IV Semiarid-mountain	In desert and high steppe regions as well as mountainous regions above 3,000 ft.	240 to 300	0 to 60	Year-round dry climate.
V Continental-Polar	In middle latitudes and polar regions. (Mediterranean climates are included because of the low range resulting from characteristic dry summers.)	300 to 340	0 to 30	Moderate or low annual mean temperatures.
VI Isothermal-equatorial	Tropical stations at low elevations between 20° N and 20° S lat., almost exclusively along seacoasts or on islands.	340 to 400	0 to 30	Monotonous rainy climates.

For a given classification of refractive-index climate, diverse meteorological climates and geographical regions may be represented. Note, for example, that type V of table 3 includes stations from Mediterranean and marine as well as polar climates. Mediterranean stations in this category fail to attain a high range because of the characteristic dryness of the subtropical high-pressure pattern that is generally found in this area during the summer months. Polar and marine climates in this group maintain a low range due to suppressed humidity effects as a result of low to moderate year-around average temperatures.

Annual trends of N_s for stations typical of each climatic division are shown by figure 14.

Yet another facet of the climate is the year-to-year variation of the monthly mean value of N_s. Five consecutive years of monthly means were prepared for each of the six typical stations whose annual cycles are shown in figure 14. Then, for each month, the absolute value of the difference between consecutive years was obtained. These values were then averaged for all months and are listed in the second column of table 4.

Another measure of the variation of monthly mean values of N_s is obtained by differencing the maximum and minimum values occurring for a given month during the 5-yr period. These differences are also given in table 4 for the months of February, May, August, and November.

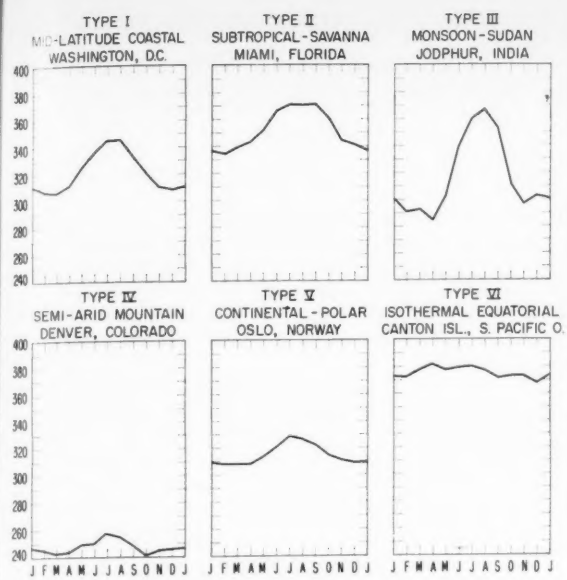


FIGURE 14. *Representative annual cycles of N_s for the major climatic types.*

TABLE 4. *Year-to-year differences of monthly mean N_s*

Climatic type	Difference between monthly means in successive years for the same month, averaged for all seasons over a 5-yr period	Maximum differences between monthly means over a 5-yr period			
		Feb.	May	Aug.	Nov.
I	5.7	6.0	16.5	17.0	7.0
II	5.4	8.5	6.5	8.0	11.0
III	8.9	16.5	14.5	20.0	6.5
IV	5.4	10.5	11.5	13.5	6.5
V	4.7	5.5	11.5	10.0	5.0
VI	7.1	9.5	25.5	8.5	8.5

5. Applications

The communications engineer usually has available a small amount of measured field strength data from limited tests of a particular system. He must then estimate the expected signal level or practical range of that system, or other systems, for other times of the year, other years and in other areas. The variation of signal level from month to month and climate to climate can be explained, in part, by its observed correlation with N_s.

Pickard and Stetson [8, 9] were among the first to note the correlation of N_s and received field strengths. The correlation of N_s and field strength over a particular path has been studied quantitatively [10, 11] and found to be highest (correlation coefficients of 0.8 to 0.95) when the variables are averaged over periods of a week to a month. This latter study has shown that the regression coefficient (decibel change in field strength per unit change in N_s) varies diurnally from 0.14 db in the afternoon hours to 0.24 db per unit change of N_s in the early morning hours. This correlation is so sufficiently consistent that Gray [12] and Norton [13] have utilized it in their recent prediction methods of transmission loss in a band from 100 to 50,000 Mc. In addition, the coefficient 0.2 db per unit change in N_s has been tentatively adopted by CCIR Study Group V in their revision of the 30- to 300-Mc tropospheric-wave propagation curves to account for the geographic and seasonal variations of field strengths. The estimates of field strength variations attributed to N_s given below are based upon the CCIR coefficient.

If one assumes, for comparison only, that the worldwide average value of N_s is 330 and that one is able to estimate the field strength level of a particular communications system at a given distance and for $N_s=330$, then the above correlations would indicate that the climatic variations of fields given in table 5 might be expected.

TABLE 5.—*Climatic variation of hypothetical communications system relative to predicted value for $N_s=330$, assuming a 0.2-db variation per unit change in N_s*

Climatic type [a]	Expected yearly mean field strength level relative to $N_s=330$	Expected annual range on the above assumption
	db	*db*
I	−6 to +4	6 to 12
II	+4 to +14	6 to 12
III	−10 to +14	12 to 20
IV	−18 to −6	0 to 12
V	−6 to +2	0 to 6
VI	+2 to +14	0 to 6

[a] Climatic types are the same as those in table 3.

The data of table 5 indicate, for example, that identically equipped tropospheric communications systems could display as much as a 32-db difference in mean signal-strength level due to the climatic difference of say, Denver, Colorado and the tropics. Further one might expect the monthly mean field strength of this hypothetical system to vary throughout the year from less than 12 db in the high plains near Denver to as much as 20 db in the African Sudan.

Under this same assumption, figures 10 and 11 allow the communications engineer to estimate the expected maximum and minimum monthly mean field strength expected throughout the year.

The year-to-year variations of the monthly mean N_s listed in table 4 indicate that the monthly mean of field strength for a particular month may differ in successive years by as little as 1.0 db for climatic category V in November or as much as 5.1 db for category VI in May.

Another application of these worldwide charts is to aid in estimating the refraction of radio waves. The most convenient method of accounting for the effects of atmospheric refraction is by means of the effective-earth's-radius concept of Schelling, Burrows, and Ferrel [14]. The effective earth's radius, a_e, is determined from

$$a_e=\left(\frac{1}{1+\frac{a}{n}\frac{dn}{dh}}\right)a, \qquad (6)$$

where a is the true radius of the earth, n is the refractive index, and dn/dh is the initial n gradient with respect to height. A great simplification of propagation calculations is accomplished by assuming that dn/dh is a constant, thus allowing radio rays to be drawn as straight rays over a fictitious earth of radius a_e rather than curved rays over the true earth of radius a. This simplification allows, for example, the distance to the radio horizon, d, of a radio ray leaving an antenna of height, h, to be calculated from

$$d=\sqrt{2a_eh}. \tag{7}$$

One notes, however, that the determination of a_e involves dn/dh as well as n and that our N_0 charts allow only an estimation of n. This disparity may be resolved by utilizing the observation that N_s is highly correlated with the value of N at 1 km above the surface. The difference between N_s and N at 1 km is denoted ΔN. It has been noted [15] that the correlation coefficient between ln $|\Delta N|$ and N_s is 0.926 for 888 sets of data from 45 U.S. weather stations representing many diverse climates. The regression equation

$$-\overline{\Delta N}=7.32\ \exp\{0.005577\ \overline{N_s}\} \tag{8}$$

results when both variables are averaged over 6 to 8 yr of record. Approximating dn/dh in (6) by ΔN we may determine that the radio horizon distance of an antenna located 150 m above the earth would vary from 48 km when $N_s=200$ to 59 km when $N_s=400$. Yet another application of the N_s charts is to the exponential models of the decrease of refractive index with height which have been recently proposed [15, 16]. These models are completely specified by N_s and may be used to account for seasonal and geographic variations of such refraction effects as radar range and elevation angle errors.

6. Critical Appraisal of Results

The world maps presented above were based upon data from 306 weather stations. This number of stations appears to be consistent with the scale of map used. The map scale is so small, however, that only large climatic differences can be expected to be discerned. For the climate of any given area one should refer to detailed studies of N such as those currently in preparation for the United States at the Central Radio Propagation Laboratory.

The accuracy of the present charts may be assessed from the charts of maximum range, R, of monthly means as given by figures 12 and 13. The standard deviation of the individual monthly means may be estimated from [17], 0.43 R, where the coefficient 0.43 is appropriate for five individual observations. Since, in general, $R\leq 20$ N units, then 0.43 $R\leq 9N$ units, although this standard deviation may be as large as 26 N units for the month of February in Australia and 17 N units in the southwest of the United States during August, or in the African Sudan during February.

Further, the standard error of estimating a five year mean from five individual monthly values is determined from

$$\frac{0.43\,R}{\sqrt{n}},$$

where n for our case is 5 and thus the error of the 5 yr mean would be 0.192 R. Remembering that $R\leq 20$ N units and assuming perfect skill in drawing the contours, one would expect the standard error of estimate to be less than 4 N units. This standard error can be as large as 12 N units in Australia where $R=60$ N units.

The value of N_s at each of the 20 test stations of section 2 was estimated from the N_0 contours with an rms error of 5 N units which is consistent with the standard error of estimate obtained from the range charts. In the large areas of sparse data, such as the oceans and Russia, this uncertainty rises to about 10 N units and thus the contours in these regions are dashed.

At the time the present study was initiated it was felt that N_s should be reduced to sea level by at least the dry term correction factor as in eq (5). The absence of published work on models of N structure in the free atmosphere encouraged the decision to rest on prudence and adopt this dry term correction factor. Since that time several effective exponential models of the free atmosphere have been demonstrated [15, 16] and it now appears that it might be better in future work to use the slightly larger exponential coefficient which corresponds to the decay of N in the free atmosphere. If the free atmosphere decay were adopted, then the range of N_0 values on the world maps would be reduced from 115 to 110 N units, i.e., less than 5 percent. Since this reduction in range is an order of magnitude less than the original reduction accomplished by the use of N_0, it appears that the basic advantage of the adoption of the concept of a reduced value has been realized with the initial correction. To produce a significantly large reduction in the range of the present map contours it would appear that the seasonal and diurnal variation of the exponential coefficient would have to be considered; a process which appears at present to be unduly complex.

7. Conclusions

With the above critical appraisal in mind, the salient conclusions of the present study are:

(a) The radio refractive index varies in a systematic fashion with climate, and different climates may be indentified by the range and mean values of the refractive index.

(b) It is 4 or 5 times more accurate to estimate the station value of the index from charts of the reduced-to-sea-level index than from charts of the station value. This improved accuracy results from using a method that allows height dependence to be accurately taken into account.

(c) Identically equipped tropospheric communications might be expected to vary as much as 30 db in monthly mean signal level in different climatic regions, and the annual range of monthly mean field strength could be as high as 20 db in the Sudan of Africa and as low as 0 to 6 db in the high plains of the western United States.

The authors express their appreciation to Mrs. G. E. Richmond, Mrs. P. C. Whittaker, and Mrs. M. F. Fischer for their aid in the preparation of the basic data, to A. M. Ozanich for cartographic assistance, and to Miss G. Hoffmire for preparation of the manuscript.

8. References

[1] E. K. Smith, Jr. and S. Weintraub, The constants in the equation for atmospheric refractive index at radio frequencies, Proc. IRE **41,** 1035 (1953).
[2] IMO list of Resolutions of the Conference of Directors, Washington, 22d Sept.–11th Oct. 1947, Lausanne, 1948, Resolution No. 71 [CCI Toronto 1947: XIII].
[3] World atlas of sea surface temperatures, Hydrographic Office Publication No. 225, 2d Ed. 1944, Washington, D.C.
[4] Atlas of climatic charts of the oceans, U.S. Weather Bureau Publication No. 1247 (Washington, D.C., 1938).
[5] W. E. Johnson, An analogue computer for the solution of the radio refractive-index equation, J. Research NBS **51,** 335 (1952) RP2462.
[6] B. Haurwitz, Dynamic meteorology, p. 11–12 (McGraw-Hill Book Co., New York, N.Y., 1941).
[7] Smithsonian Meteorological Tables, Table 63, Sixth Revised Ed. (Washington, D.C., 1951).
[8] G. W. Pickard and H. T. Stetson, Comparison of tropospheric reception, J. Atmospheric and Terrest. Phys. **1,** 32 (1950).
[9] G. W. Pickard and H. T. Stetson, Comparison of tropospheric reception at 44.1 Mc with 92.1 Mc over the 167-mile path of Alpine, N.J. to Needham, Mass., Proc. IRE **38,** 1450 (1950).
[10] B. R. Bean, Some meteorological effects on scattered radio waves, IRE Trans. **PGCS–4** 32 (1956).
[11] M. Onoe, M. Hirai, and S. Niwa, Results of experiment of long-distance overland propagation of ultra-short waves, J. Radio Research Labs. (Japan) **5,** 79 (1958).
[12] R. E. Gray, The refractive index of the atmosphere as a factor in tropospheric propagation far beyond the horizon, IRE National Conv. Record, **Pt 1,** 3 (1957).
[13] K. A. Norton, Point-to-point radio relaying via the scatter mode of trophospheric propagation, IRE Trans. **PGCS–4,** 39 (1956).
[14] J. C. Schelling, C. R. Burrows, and E. B. Ferrell, Ultra-Short-Wave Propagation, Proc. IRE **21,** 427 (1933).
[15] B. R. Bean and G. D. Thayer, On models of the atmospheric refractive index, Proc. IRE **47,** 740 (1959).
[16] L. J. Anderson, Tropospheric bending of radio waves, Trans. Am. Geophys. Union **39** 208 (1958).
[17] G. W. Snedecor, Statistical methods, 4th ed., p. 97–98 (Iowa State College Press, Ames, Iowa, 1946).

BOULDER, COLO. (Paper 63D3–22)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Path Antenna Gain in an Exponential Atmosphere

W. J. Hartman and R. E. Wilkerson

(May 7, 1959)

The problem of determining path antenna gain is treated here in greater detail than previously [1, 2, 3].[1] The method used here takes into account for the first time the exponential decrease of the gradient of refractive index with height, and a scattering cross section inversely proportional to the fifth power of the scattering angle. Results are given for all combinations of beamwidths and path geometry, assuming that symmetrical beams are used on both ends of the path and that atmospheric turbulence is isotropic. The result appears as a function of both of the beamwidths, in addition to other parameters, and thus the loss in gain cannot be determined independently for the transmitting and receiving antennas. The values of the loss in gain are generally lower than the previous estimates for which a comparison is possible.

1. Introduction

Theories of the forward scatter of radio waves in the troposphere or stratosphere usually differ only in their assumptions about the radio wave scattering cross section and pertinent meteorological parameters. With any theory, the total effective gain of highly directive antennas is less over a scatter circuit than it would be in free space. The "path antenna gain" is less than or equal to the sum of the free space gains G_t and G_r of the transmitting and receiving antennas. Expressed in decibels above the gain of an isotropic radiation [1],

$$G_p \leq G_t + G_r. \tag{1-1}$$

Booker and de Bettencourt [2], Norton, Rice, and Vogler [1], and more recently Staras [3] have estimated the loss in antenna gain G_L, where G_L in decibels is given by

$$G_L = G_t + G_r - G_p, \tag{1-2}$$

for an atmosphere in which the variance of the gradient of the radio refractive index is either constant or varies parabolically with height. This paper estimates G_L with a model atmosphere in which the refractive index decreases exponentially with height. G_L is evaluated using antenna radiation patterns which are the same horizontally and vertically. To the authors' knowledge, G_L has not been evaluated using antenna patterns which are not symmetric about the beam axis. Figures 6 and 7 show graphs of G_L for various values of the parameters, and section 4 summarizes important definitions and gives a simplified method for predicting G_L. The reader who desires only an estimate of G_L may skip to section 4.

2. Antenna Gain in Free Space

The gain function $f(\theta, \phi)$ is the ratio of the power radiated in the direction (θ, ϕ) per unit solid angle to the average power radiated per unit solid angle. This expresses the increase in power radiated in a given direction by the antenna over that from an isotropic radiator emitting the same total power. It is convenient to express the gain pattern of an antenna and its free space gain in terms of the electric field distribution $F(\xi, \eta)$ over the antenna aperture. Figure 1 shows a coordinate system with an antenna aperture in the x, y plane and centered at the origin.

[1] Figures in brackets indicate the literature references at the end of this paper.

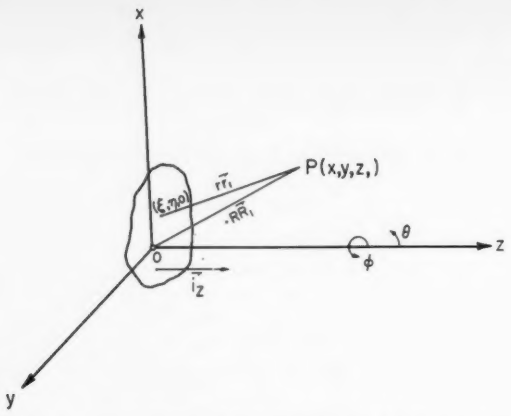


FIGURE 1. *Geometry for determining the gain of an antenna.*

The point $P(x, y, z)$ is in free space in the Fraunhofer or far zone. Using this notation, the power radiated per unit solid angle in a given direction by an antenna having constant phase distribution over the aperture is approximately [4]

$$P(\theta,\phi)=\frac{\sqrt{\epsilon/\mu}}{2\lambda^2}\left|\int_A F(\xi,\eta)\exp\left[jk\sin\theta(\xi\cos\phi+\eta\sin\phi)\right]d\xi d\eta\right|^2, \tag{2–1}$$

where $\sqrt{\mu/\epsilon}$ is the impedance of the propagating medium, λ is the free space radio wavelength, $k=2\pi/\lambda$, and $d\xi d\eta$ is a surface element of the aperture. The average power radiated per steradian in terms of the total power P_a across the aperture is

$$P_a/4\pi=\frac{\sqrt{\epsilon/\mu}}{8\pi}\int_A|F(\xi,\eta)|^2d\xi d\eta, \tag{2–2}$$

and the ratio of (2–1) to (2–2) is the gain function.

$$f(\theta,\phi)=4\pi P(\theta,\phi)/P_a. \tag{2–3}$$

If the aperture illumination $F(\xi,\eta)$ is symmetrical about the origin, the maximum value $f_{max}(\theta,\phi)$ of (2–3) corresponds to the direction $\theta=0$ along the antenna beam axis, the z-axis in figure 1,

$$f_{max}(\theta,\phi)=f(0,\phi)=A_e/(\lambda^2/4\pi). \tag{2–4}$$

The factor A_e, called the "effective absorbing area" of an antenna, is given in terms of the distribution function as

$$A_e=\left|\int_A F(\xi,\eta)d\xi d\eta\right|^2\Big/\int_A|F(\xi,\eta)|^2d\xi d\eta. \tag{2–5}$$

For an isotropic radiator, $A_e=\lambda^2/4\pi$, as follows from (2–1) and (2–2). For any antenna with a uniformly illuminated aperture ($F(\xi,\eta)$ constant), the effective area equals the actual area. as may readily be seen from (2–5). For the more usual 10 db tapered illumination of the aperture of a parabolic dish antenna, the effective area A_e is about 56 percent of the actual area.

The free-space gain of an antenna in decibels above the gain of an isotropic radiator is

$$G=10\log_{10}f_{max}(\theta,\phi)=10\log_{10}[A_e/(\lambda^2/4\pi)]\text{db}. \tag{2–6}$$

Define the normalized gain function, $g_m(\theta,\phi)$:

$$g_m(\theta,\phi)=f(\theta,\phi)/f_{max}(\theta,\phi). \tag{2–7}$$

Figure 2 shows the normalized gain function $g_m(\theta,\phi)$ versus θ for a uniformly illuminated circular aperture of 100 wavelengths diameter. Figure 2 would be the same for any value of ϕ. Figure 2 also shows a "square" antenna power pattern, where $g_m(\theta,0)$ maintains the value unity for $0 \leq \theta \leq \delta$ and is zero when θ exceeds δ.

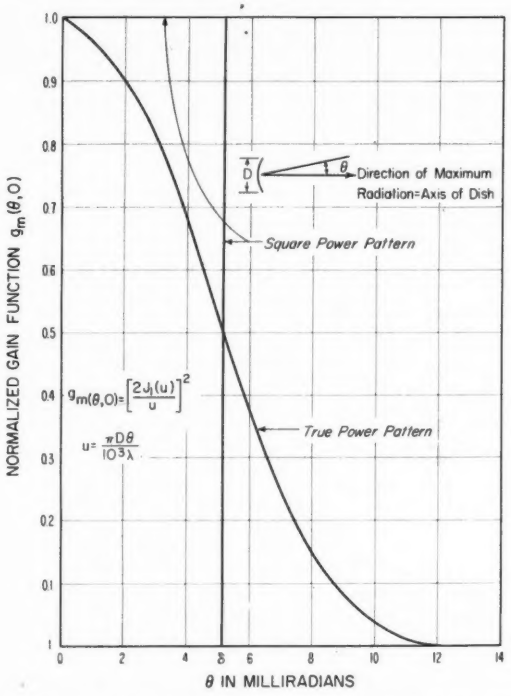


FIGURE 2. *Computed polar pattern (without side lobes) for a parabolic dish (circular aperture uniformly illuminated) with* $D/\lambda = 100$

The "square" power pattern will be used throughout the calculations in this paper.

For practical purposes the gain is often specified in terms of the half-power beamwidth, 2δ. For an antenna with a symmetrical distribution of field intensity over the aperture one obtains [4]

$$2\delta \approx \lambda\sqrt{\mu_0/\mu_2}, \tag{2-8}$$

where μ_i is the ith moment of the function $F(x)$ which gives the aperture field in terms of $x=\sqrt{\xi^2+\eta^2}$. Combining (2-1), (2-7), and (2-8).

$$f(0,\phi) \approx (\mu_0/\mu_2)/(2\delta)^2 \tag{2-9a}$$

or

$$f(0,\phi) = C_1/(2\delta)^2, \tag{2-9b}$$

where C_1 is a constant of proportionality depending on the characteristics of the antenna used. As examples, with the beamwidth 2δ expressed in radians, $C_1=7.2$ for a half-wave dipole, and $C_1=8.2$ for a parabolic dish with the edge illumination 10 db down from the center illumination. This constant C_1 can be determined for some types of antennas using tables in the book, *Reference Data for Radio Engineers* [5].

The half-power beamwidth in figure 2 is 2δ, where δ is the value of θ corresponding to $g_m(\theta) = 0.5$ For the square pattern, assuming that δ is small, the power is radiated through $\pi\delta^2$ steradians if the aperture A is circular, and through $(2\delta)^2$ steradians if A is square. The power gain $f(0,\phi)$ is then $4\pi/\pi\delta^2$ for the circular aperture or $4\pi/(2\delta)^2$ for the square aperture, that is, $C_1=16$ or $C_1=4\pi$.

G_L, as determined by calculations in this report, depends directly on the semibeamwidths δ_t and δ_r, and to a lesser extent on the shapes of the associated power patterns. Subscripts t and r in this paper refer to transmitting and receiving antennas, respectively.

3. Scatter Integral

If we denote by K the received power relative to free space of a given scatter propagation circuit with arbitrary antennas, and by K_0 the received power relative to free space of the same circuit using isotropic antennas, then the loss in gain in decibels may be defined as

$$G_L = -10 \log_{10} \frac{K}{K_0} \tag{3-1}$$

and the total path antenna gain is

$$Gp = G_t + G_r - G_L. \tag{1-2}$$

The loss in gain G_L depends on the plane wave power patterns of the two antennas and upon the path parameters in such a way that the effect of one antenna cannot be evaluated independently of the power pattern of the other.

In this report, it will be assumed that the antennas are sufficiently high so that ground reflection effectively doubles the power incident on the scattering volume over what it would be in free space and again doubles the power at the receiver. In general, K is given by the volume integral

$$K = d^2 \int_V \frac{g_t g_r \sigma dv}{(R R_0)^2}, \tag{3-2}$$

where d is the straight line distance between transmitter and receiver, R and R_0 are distances from transmitter and receiver, respectively, to a scattering element dv, and g_t and g_r are $2g_{mt}$ and $2g_{mr}$, that is, double the power patterns defined by (2–7). The scattering cross section σ gives the fractional part of the incident power scattered in a given direction (at an angle, $\alpha+\beta$ from the forward direction) per unit volume per unit solid angle. Refer to figure 3, which shows

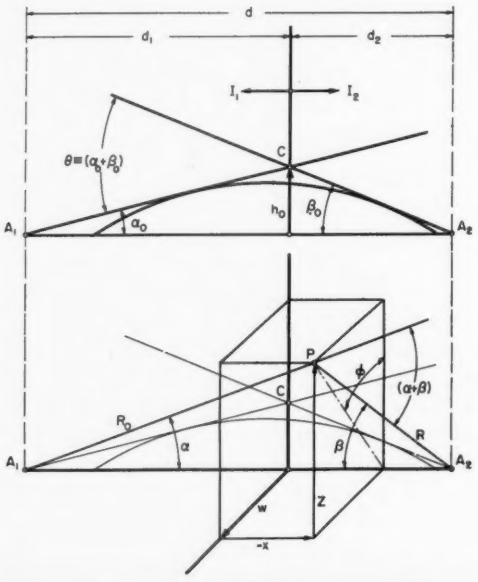


FIGURE 3. *Geometry for the scattering integral.*

some of the geometry appropriate for the evaluation of (3–2). In this figure and throughout the remainder of this paper, θ is the angular distance, or the minimum scattering angle. A detailed discussion of the derivation of the form for σ used here may be found in a recent paper [6]. The following result is obtained:

$$\sigma=\frac{3\lambda<(\Delta n)^2>}{4\pi l_0^{\,2}}\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^{-5}, \tag{3–3}$$

where $\alpha+\beta$ is the angle between the directions R and R_0 as shown in figure 3, λ is the wavelength, $<(\Delta n)^2>$ is the variance of refractive index fluctuations in the scattering medium, and l_0 is a "scale of turbulence" appropriate for the scattering medium, assuming homogeneous and isotropic turbulence. Following Booker and Gordon [7], Abbott [8] derived in a substantially different fashion the following expression for $<(\Delta n)^2>/l_0^2$:

$$<(\Delta n)^2>/l_0^2=S' \exp(-2\gamma h)=S\exp(-2\gamma z) \tag{3–4}$$

where S' is the value of this parameter at sea level, $h\cong z-h_0/2$, $S\cong S'\exp(\gamma h_0)$, and $1/\gamma$ is the scale height of the atmosphere, here assumed constant.

Using the coordinate system shown in figure 3, (3–2) becomes

$$K=\frac{3\lambda S d^2}{\pi}\int_{x_1}^{x_2}dx\int_{z_1}^{z_2}dz\int_{-\infty}^{\infty}dw\frac{g_{mt}g_{mr}\exp(-2\gamma z)}{\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^5(RR_0)^2}. \tag{3–5}$$

Using the approximation

$$\frac{1}{\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^5(RR_0)^2}\cong\frac{(d_1-x)^3(d_2+x)^3}{d^5(w^2+z^2)^{5/2}}, \tag{3–6}$$

which is justified in appendix 1, and assuming that the effects of back scatter are negligible so that we may limit our x integration to the region between the transmitter and receiver, and limiting z to the region above the horizon rays, we obtain

$$K=\frac{3\lambda}{4\pi d^3}\left\{\int_0^{d_1}dx\int_{z_{1m}}^{\infty}dz\int_{-\infty}^{\infty}dwf(v)+\int_{-d_2}^{0}dx\int_{z_{2m}}^{\infty}dz\int_{-\infty}^{\infty}dwf(v)\right\}, \tag{3–7}$$

where

$$f(v)=g_{mt}g_{mr}\frac{(d_1-x)^3(d_2+x)^3Se^{-2\gamma z}}{(w^2+z^2)^{5/2}} \tag{3–8}$$

and

$$z_{1m}\cong\beta_0(d_2+x) \tag{3–9}$$

$$z_{2m}\cong\alpha_0(d_1-x) \tag{3–10}$$

We can express K in terms of four parameters, g_{mt}, g_{mr},

$$s=\frac{\alpha_0}{\beta_0} \tag{3–11}$$

and

$$\eta=\gamma d\theta, \tag{3–12}$$

where

$$\theta=\alpha_0+\beta_0, \tag{3–13}$$

Thus

$$K=K(s,\ \eta,\ g_t,\ g_r). \tag{3–14}$$

For isotropic antennas $g_{mt}=g_{mr}=1$, and thus $g_t=g_r=2$, including the effects of ground reflection as explained previously. K then becomes

$$K(s,\eta,2,2)\equiv K_0=\frac{3\lambda S}{\pi d^3}\left\{\int_0^{d_1}dx\int_{z_{1m}}^{\infty}dz\int_{-\infty}^{\infty}dw\,\frac{(d_1-x)^3(d_2+x)^3e^{-2\gamma z}}{(w^2+z^2)^{5/2}}\right.$$

$$\left.+\int_{-d_2}^{0}dx\int_{z_{2m}}^{\infty}dz\int_{-\infty}^{\infty}dw\,\frac{(d_1-x)^3(d_2+x)^3e^{-2\gamma z}}{(w^2+z^2)^{5/2}}\right\}. \tag{3-15}$$

(3–15) is evaluated exactly in appendix 2 with the result given by (7–14). K (3–7) was evaluated by imposing geometric limits on the integrals and using Taylor series expansions. The limits vary as $s\mu<1$, $s\mu=1$, or $s\mu>1$ where

$$\mu=\frac{\delta_r}{\delta_t}. \tag{3-16}$$

Figure 4 is a composite drawing and figure 5 is a drawing showing the xw and xh planes for $s\mu<1$. This leads to five integrals corresponding to the regions shown in figure 5 since it was found necessary to evaluate the integrals corresponding to the regions I_{42} and I_{43} as one integral, and then subtract the integral for I_{43}. The integral K was also evaluated numerically with $<(\Delta n)^2>/l_0^2$ proportional to z^{-2} and z^{-4}, using the antenna power pattern $[2J_1(u)/u]^2$ of figure 2 (without side lobes), and the value obtained for G_L was found to be always within one and a half decibels of the value obtained using a "square" pattern with the same half-power beamwidth and height dependence.

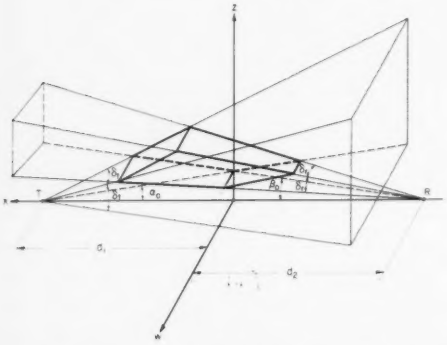


FIGURE 4. *Scattering volume with "square" antenna patterns.*

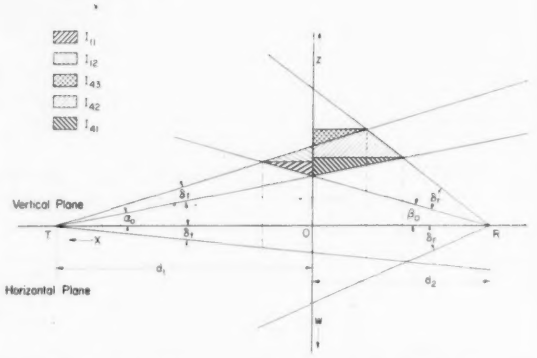


FIGURE 5. *Shape of scattering volume in vertical and horizontal planes, $s\mu>1$.*

Effects of anisotropic turbulence have not been evaluated in this paper, but it can be shown that anisotropic turbulence arising from scale length changes has no effect on K_0 [isotropic antennas] except to add a multiplicative factor of r, where

$$r=\frac{l_v}{l_w}=\frac{\text{scale of turbulence in vertical dimension}}{\text{scale of turbulence in horizontal dimension}}. \tag{3-17}$$

However, it has been shown [6, 9] that anisotropic turbulence produces the following effect in the integrand, (3–8) of the scatter integral (3–7)

$$f_A(v)=g_{mt}g_{mr}\frac{(d_1-x)^3(d_2+x)^3Se^{-2\gamma z}}{(w^2l_w^{\,2}+z^2l_v^{\,2})^{5/2}}$$

$$=\frac{g_{mt}g_{mr}}{l_w^{\,5}}\frac{(d_1-x)^3(d_2+x)^3Se^{-2\gamma z}}{\left(\left(\frac{w}{r}\right)^2+z^2\right)^{5/2}}, \tag{3-18}$$

where $f_A(v)$ is $f(v)$ with the effects of anisotropy included. Using the square pattern as was done in this paper, with geometric limits acting as a cutoff for the antenna beamwidths, it will be shown that, at least qualitatively, anisotropy will have some effect when using high gain antennas. As can be seen using (3–18), only the w integration is affected. A typical w integral using the geometric limits for anisotropic turbulence would be of the form

$$\int_0^{\delta(d_2+x)} \frac{dw}{\left[\left(\frac{w}{r}\right)^2+z^2\right]^{5/2}}. \tag{3–19}$$

Let $\frac{w}{r}=v$. Then (3–19) becomes

$$r\int_0^{\frac{\delta}{r}(d_2+x)} \frac{dv}{(v^2+z^2)^{5/2}} \tag{3–20}$$

This integrand is the same as in the case of isotropic turbulence, but the range of integration is larger, since $r<1$. Thus, the values of G_L obtained here would be lower if the effects of anisotropy were included. This effect is the same as if antennas with larger beamwidths in the horizontal dimension than in the vertical dimension were used in an isotropic atmosphere.

4. Results

The integrals in the preceding section were calculated for numerous combinations of the parameters. Typical curves which resulted from these calculations are plotted in figure 6. The curves in figure 6 corresponding to path symmetry ($s=1$), and identical antennas ($\mu=1$), are shown again in figure 7 using a different coordinate system which facilitates interpolation. Because of the number of parameters involved it was found impractical to show all of the curves of the type shown in figure 6, and instead, an empirical formula was derived, using these curves, which gives good accuracy using only two graphs and a few simple calculations. The following discussion defines the necessary parameters and shows how to use the empirical curves.

The basic parameters which depend on the path geometry are shown in figure 3 and are as follows:

$$\alpha_0, \quad \beta_0, \quad d, \tag{4–1}$$

from which may be calculated

$$s=\frac{\alpha_0}{\beta_0}, \tag{4–2a}$$

$$\theta=\alpha_0+\beta_0 \tag{4–2b}$$

and

$$h_0=\frac{s}{(1+s)^2}d\theta. \tag{4–2c}$$

Combining these with the rate of decay of the variance of the refractive index with height, γ, one obtains

$$\eta=\gamma d\theta \tag{4–3a}$$

and

$$v=2\gamma h_0=\frac{2\eta s}{(1+s)^2}. \tag{4–3b}$$

For the purpose of illustration,[2] $\gamma=0.23$/mile will be used here.

[2] For additional values of γ, see reference [10].

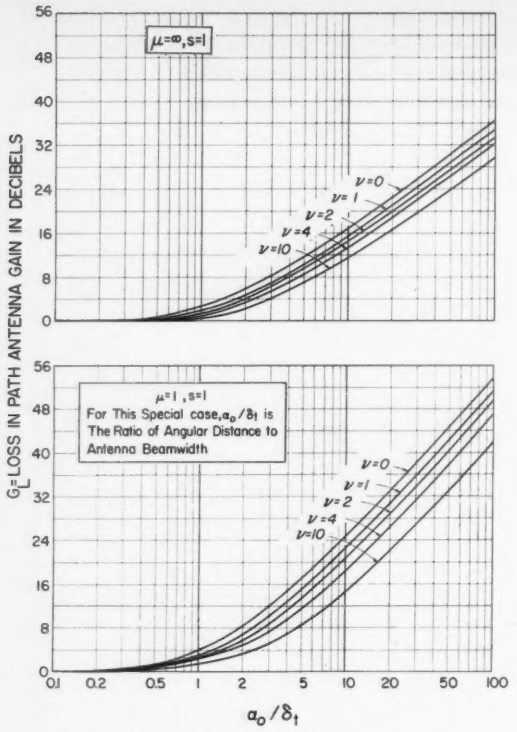


FIGURE 6. *Loss in path antenna gain.*

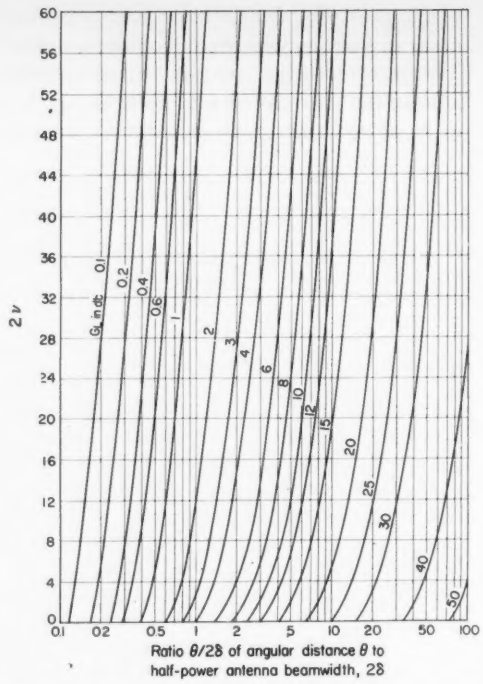


FIGURE 7. *Loss in antenna gain,* G_L *assuming equal free space gains* G_t *and* G_r *at the terminals of a symmetrical path* $\delta_t = \delta_r = \delta$, $s=1$.

The parameters which depend on the antenna characteristics are the half power semibeamwidths, δ_t and δ_r. These are used to calculate

$$\mu = \delta_r/\delta_t. \tag{4-4}$$

At the end of this section, formulas are given for calculating δ from the free space gain of the antenna or from the frequency and area of the antenna aperture.

In the empirical formula, the two parameters s and μ, are combined in the single parameter $s\mu$, and two separate cases arise.

Case I, $s\mu \geq 1$.

Calculate

$$n = \alpha_0/\delta_t. \tag{4-5}$$

Read $f(\nu)$ from figure 8 and calculate

$$\hat{n} = (n + 0.03\nu)/f(\nu). \tag{4-6}$$

Read G_L from figure 9 for the appropriate value of $s\mu$. Linear interpolation is possible using the formula

$$G_L = G_L' + \frac{\frac{1}{s\mu} - \frac{1}{(s\mu)'}}{\frac{1}{(s\mu)''} - \frac{1}{(s\mu)'}} (G_L'' - G_L'), \tag{4-7}$$

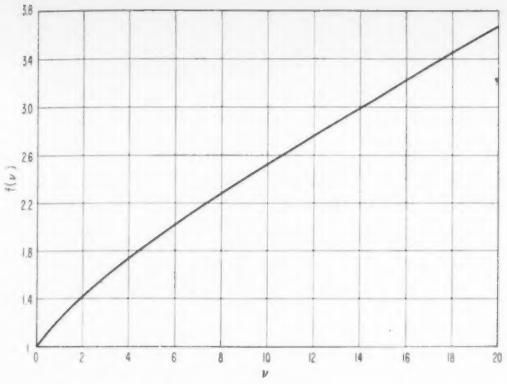


FIGURE 8. *The contraction factor* $f(\nu)$.

Supplementary Values.

ν	$f(\nu)$	ν	$f(\nu)$
1	1.22	5	1.88
2	1.41	15	3.10

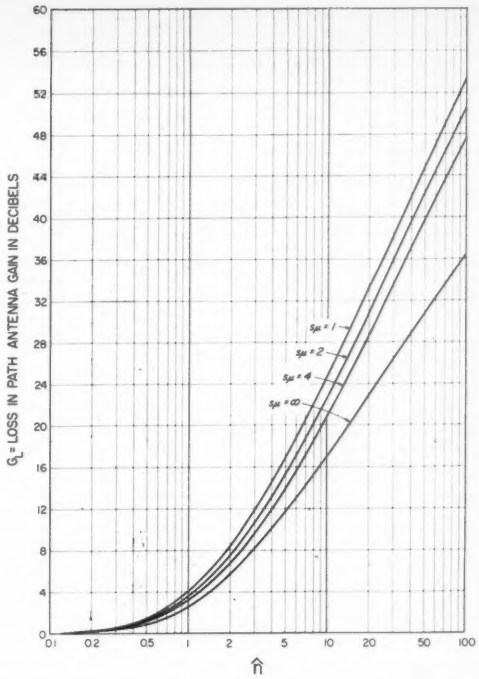


FIGURE 9. *Loss in path antenna gain versus* $\hat{n}$ *and the product* $s\mu$ *for* $\nu=0$.

where G_L' and G_L'' correspond to the graphically obtained values for $(s\mu)'$ and $(s\mu)''$ less than $s\mu$ and greater than $s\mu$, respectively.

Case II, $s\mu<1$.

Calculate

$$n=\beta_0/\delta_r. \tag{4–8}$$

Read $f(\nu)$ from figure 8 and calculate

$$\hat{n}=(n+0.03\nu)/f(\nu). \tag{4–9}$$

Read G_L from figure 9 on the curve corresponding to the value of $1/s\mu$. Use formula (4–7) for interpolating replacing $(s\mu)$ by $(1/s\mu)$, $(s\mu)'$/ by $(1/s\mu)'$ and $(s\mu)''$ by $(1/s\mu)''$.

If the gain G in decibels is known, the half-power semibeamwidth may be calculated using the following formulas which may be derived from the formulas in section 2, and are given here for a parabolic dish of diameter D feet at a frequency f_{mc} megacycles per second. G is given by

$$G=10\log_{10}[0.56\pi^2(D/\lambda)^2] \tag{4–10a}$$

$$=20\log_{10}D_{ft}+20\log_{10}f_{mc}-52.431 \text{ db}.$$

Also,

$$G=10\log_{10}[8.2/(2\delta)^2] \tag{4–10b}$$

with δ in radians, and thus,

$$\delta=1432\times10^{-G/20}=609\lambda/D \text{ mr}. \tag{4–10c}$$

5. Comparison with Previous Estimates

In figure 10, our estimates of G_L are plotted together with the estimates given by the previous authors [1, 2, 3]. The values from Staras' [3] paper which are shown in figure 10 correspond to the assumption of isotropic turbulence. For this comparison, it was necessary to use only the curves corresponding to the assumptions of path symmetry and identical antennas on each end of the path since some of the previous estimates are available only for these cases. In addition, our assumption that the gradient of refractivity of the atmosphere decreases exponentially with height leads to the parameter $\nu=2\gamma h_o$ which does not appear in the other papers. For making this comparison we chose the typical value of $\nu=2$ corresponding to $\gamma\simeq 0.23$, $d\simeq 300$ mi, and $\theta\simeq 50$ mr. Our values of G_L in figure 10 are lower than the other estimates, and

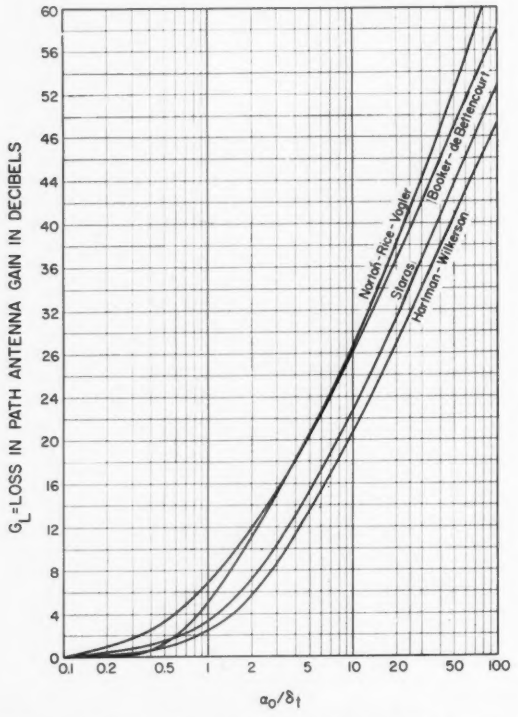


FIGURE 10. *A comparison of loss in path antenna gain obtained from different scatter theories for a symmetrical path for* d=300 *miles,* $\mu=\delta_r/\delta_t=1$, $s=\alpha_o/\beta_o=1$ *(antennas on the surface).*

as can be seen in figure 6, they would be even lower if a larger value of ν were used or if an allowance were made for some degree of anisotropy. The difference between our curve and the others is primarily the result of the use of the exponential model atmosphere which is supported by ample data [10] and a scattering cross section that is proportional to $[2 \sin \{(\alpha+\beta)/2\}]^{-5}$.

The authors acknowledge the assistance of P. L. Rice and H. K. Hansen in the preparation of this paper.

6. Appendix 1

The approximation

$$\frac{1}{\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^5(RR_0)^2}\cong\frac{(d_1-x)^3(d_2+x)^3}{d^5(w^2+z^2)^{5/2}} \tag{3-6}$$

is usually derived by assuming that $\alpha+\beta$ is small [1]. The following method is used to obtain this approximation, because it has the advantage that the error is evaluated. Let

$$P=\frac{R+R_0+d}{2}. \tag{6-1}$$

Then, using a well known formula from geometry,

$$\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^{-5}=2^{-5}\left\{\left(\frac{(P-R_0)(P-d)}{R_0d}\right)^{1/2}\left(\frac{P(P-R_0)}{Rd}\right)^{1/2}+\left(\frac{P(P-R)}{R_0d}\right)^{1/2}\left(\frac{(P-R)(P-d)}{Rd}\right)^{1/2}\right\}^{-5}=2^{-5}\left\{\frac{(R+R_0)^2-d^2}{4RR_0}\right\}^{-5/2}. \tag{6-2}$$

From the geometry

$$R_0=\sqrt{(d_1-x)^2+z^2+w^2} \tag{6-3}$$

$$R=\sqrt{(d_2+x)^2+z^2+w^2} \tag{6-4}$$

and

$$RR_0=(d_1-x)(d_2+x)\left\{1+\frac{z^2+w^2}{(d_1-x)^2}\right\}^{1/2}\left\{1+\frac{z^2+w^2}{(d_2+x)^2}\right\}^{1/2}. \tag{6-5}$$

(6–5) can be expanded if $z^2+w^2<(d_1-x)^2$ and $z^2+w^2<(d_2+x)^2$. This is generally the case for a beyond-the-horizon propagation path. Then,

$$RR_0=(d_1-x)(d_2+x)+\frac{z^2+w^2}{2}\left\{\frac{(d_2+x)^2+(d_1-x)^2}{(d_1-x)(d_2+x)}\right\}+\cdots. \tag{6-6}$$

Using (6–3), (6–4), and the first two terms of (6–6), we obtain

$$(R+R_0)^2-d^2\cong 2(z^2+w^2)\left\{1+\frac{(d_2+x)^2+(d_1-x)^2}{2(d_1-x)(d_2+x)}\right\}=\frac{(z^2+w^2)d^2}{(d_1-x)(d_2+x)}. \tag{6-7}$$

Using (6–2), (6–7), and the first term of (6–6) we obtain

$$\frac{\left[2\sin\left(\frac{\alpha+\beta}{2}\right)\right]^{-5}}{(RR_0)^2}=\frac{(RR_0)^{1/2}}{[(R+R_0)^2-d^2]^{5/2}}\cong\frac{(d_1-x)^3(d_2+x)^3}{d^5(w^2+z^2)^{5/2}}, \tag{6-8}$$

which is the same as (3–6).

7. Appendix 2

The derivation of (3–15) with $g_{mt}=1$, $g_{mr}=1$ is as follows:

$$K(s,\eta,2,2)=\frac{3\lambda S}{\pi d^3}\left\{\int_0^{d_1}dx\int_{z_{1m}}^{\infty}dz\int_{-\infty}^{\infty}dw\,\frac{(d_1-x)^3(d_2+x)^3e^{-2\gamma z}}{(w^2+z^2)^{5/2}}\right.$$

$$\left.+\int_{-d_2}^{0}dx\int_{z_{2m}}^{\infty}dz\int_{-\infty}^{\infty}dw\,\frac{(d_1-x)^3(d_2+x)^3e^{-2\gamma z}}{(w^2+z^2)^{5/2}}\right\}=K_1+K_2, \quad (7\text{-}1a)$$

where K_1 and K_2 correspond to the scattering regions to the left and right of the crossover C of radio horizon rays in figure 3.

$$z_{1m}=\beta_0(d_2+x). \quad (7\text{-}1b)$$

$$z_{2m}=\alpha_0(d_1-x). \quad (7\text{-}1c)$$

In K_1, let

$$x=kd_1 \quad (7\text{-}2)$$

and

$$z_0=\frac{\alpha_0 d_1}{z}\left(1+\frac{k}{s}\right), \quad (7\text{-}3)$$

where

$$s=\alpha_0/\beta_0=d_2/d_1.$$

In K_2 let

$$x=-kd_2 \quad (7\text{-}4)$$

and

$$z_0=\frac{\beta_0 d_2}{z}(1+sk). \quad (7\text{-}5)$$

Then, integrating with respect to w, we obtain

$$K_0=\frac{4\lambda S d}{\pi\theta^3(1+s)}\left\{\int_0^1(1-k)^3dk\int_0^1 z_0^2\exp\left[\frac{-\nu}{z_0}(1+k/s)\right]dz_0\right.$$

$$\left.+s\int_0^1(1-k)^3dk\int_0^1 z_0^2\exp\left[\frac{-\nu}{z_0}(1+sk)\right]dz_0\right\}, \quad (7\text{-}6)$$

where we have set

$$\nu=2\gamma h_0=2\gamma\alpha_0 d_1=2\gamma\beta_0 d_2. \quad (4\text{-}3)$$

Let

$$L_1=\frac{1}{1+s}\int_0^1(1-k)^3dk\int_0^1 z_0^2\exp\left[\frac{-\nu}{z_0}(1+k/s)\right]dz_0, \quad (7\text{-}7a)$$

and

$$L_2=\frac{s}{1+s}\int_0^1(1-k)^3dk\int_0^1 z_0^2\exp\left[\frac{-\nu}{z_0}(1+sk)\right]dz_0. \quad (7\text{-}7b)$$

We need only consider (7–7a) since (7–7b) can be obtained from (7–7a) by replacing s by $1/s$.

Thus, performing the integration with respect to k, we obtain:

$$L_1=\frac{1}{1+s}\int_0^1 z_0^2\exp(-\nu/z_0)\left\{\frac{sz_0}{\nu}-\frac{3(sz_0)^2}{\nu^2}+\frac{6(sz_0)^3}{\nu^3}-\frac{6(sz_0)^4}{(\nu)^4}+\frac{6(sz_0)^4\exp[-\nu/(sz_0)]}{\nu^4}\right\}dz_0. \quad (7\text{-}8)$$

Let

$$\frac{\nu}{z_0}=u, \tag{7-9}$$

and

$$\nu_2=\nu(1+1/s), \tag{7-10}$$

$$c_1=\frac{s^4+7s^3+21s^2+35s}{5}. \tag{7-11}$$

Then

$$L_1=\frac{\nu^3}{1+s}\int_\nu^\infty\left\{e^{-u}\left[\frac{s}{u^5}-\frac{3s^2}{u^6}+\frac{6s^3}{u^7}-\frac{6s^4}{u^8}\right]+\frac{6s^4e^{-u\left(1+\frac{1}{s}\right)}}{u^8}\right\}du$$

$$=\frac{\nu^3}{7(1+s)}\left\{e^{-\nu}\left[-\frac{6s^4}{\nu^7}+\frac{s^4+7s^3}{\nu^6}-\frac{s^4+7s^3+21s^2}{5\nu^5}+\frac{c_1}{24}\left(\frac{6}{\nu^4}-\frac{2}{\nu^3}+\frac{1}{\nu^2}-\frac{1}{\nu}\right)\right]-\frac{c_1}{24}\,Ei(-\nu)\right.$$

$$\left.+6s^4(1+1/s)^7\left[\frac{1}{\nu_2^7}-\frac{1}{6\nu_2^6}+\frac{1}{30\nu_2^5}-\frac{1}{120\nu_2^4}+\frac{1}{360\nu_2^3}-\frac{1}{720\nu_2^2}+\frac{1}{720\nu_2}+\frac{1}{720}\,Ei(-\nu_2)\right]\right\}. \tag{7-12}$$

where

$$-Ei(-x)=\int_x^\infty\frac{e^{-t}}{t}\,dt. \tag{7-13}$$

Thus, adding L_1 and L_2 we obtain:

$$K_0=\frac{\lambda Sd}{210\,\pi\theta^3\nu^4(s+1)s^3}\{e^{-\nu}[-720\,(s^7+1)+120\,(s^7+7s^6+7s+1)\nu-24\nu^2[(s+1)^7$$
$$-35s^3(s+1)]+(s+1)^7\nu^3(6-2\nu+\nu^2-\nu^3)]-(s+1)^7\nu^7[Ei(-\nu)-Ei(-\nu[1+1/s])$$
$$-Ei(-\nu[1+s])]+se^{-\nu(1+1/s)}[720s^6-120s^5(1+s)\nu+24s^4(1+s)^2\nu^2-6s^3(1+s)^3\nu^3$$
$$+2s^2(1+s)^4\nu^4-s(1+s)^5\nu^5+(1+s)^6\nu^6]+e^{-\nu(1+s)}[720-120(1+s)\nu+24(1+s)^2\nu^2$$
$$-6(1+s)^3\nu^3+2(1+s)^4\nu^4-(1+s)^5\nu^5+(1+s)^6\nu^6]\}. \tag{7-14}$$

Figure 11 shows $3\pi\theta^3(1+s)\,K_0/2\lambda Sd$ plotted versus $\eta/2$ for several values of the path asymmetry factor, s.

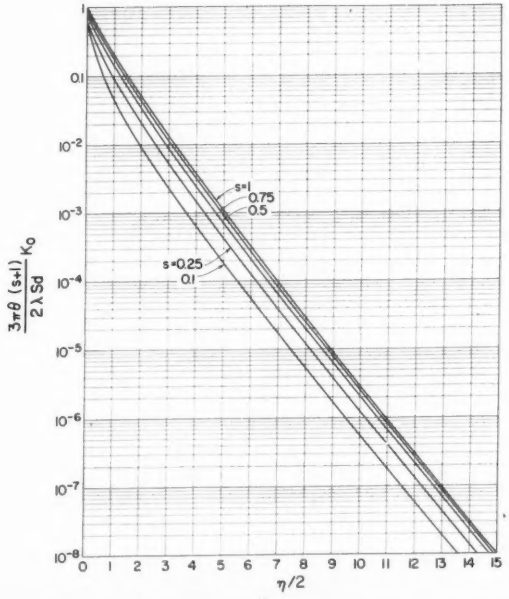


FIGURE 11. *Dependence of the scattering integral on* s *and* η.

8. References

[1] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** 1428, (1955).

[2] H. G. Booker and J. T. de Bettencourt, Theory of radio transmission by tropospheric scattering using very narrow beams, Proc. IRE **43,** 281, (1955).

[3] Harold Staras, Antenna-to-medium coupling loss, PGAP **AP–5,** 288, (1957).

[4] Samuel Silver, Microwave antenna theory and design, pp. 169–199 (McGraw-Hill, 1949).

[5] Reference data for radio engineers, International Telephone and Telegraph Corp., p. 703, 4th ed.

[6] K. A. Norton, Point-to-point relaying via the scatter mode of tropospheric propagation, Trans. IRE, Prof. Gp. on Comm. Systems **CS–4,** No. 1, 39, (1956).

[7] H. G. Booker and W. E. Gordon, The role of stratospheric scattering in radio communication, E. E. Report (Cornell University, Ithaca, New York, 1956).

[8] R. L. Abbott (private communication).

[9] H. Staras, Forward scattering of radio waves by anisotropic turbulence, Proc. IRE **43,** 1374 (1955).

[10] B. R. Bean and G. D. Thayer, Models of atmospheric radio refractive index, to be published in Proc. IRE **47,** 5, (1959). (Final pages to be set in proof.)

BOULDER, COLO.

(Paper 63D3–23)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Effect of Atmospheric Horizontal Inhomogeneity Upon Ray Tracing[1]

B. R. Bean and B. A. Cahoon

(May 7, 1959)

The tracing of radio rays is normally carried out under the assumption that the refractive index varies only in the vertical direction. Although this assumption appears to be quite reasonable in the average or climatic sense, it is seldom satisfied under actual conditions and is strongly violated by horizontal airmass changes occurring near frontal and land-sea interfaces. This latter case is investigated by tracing rays through two instances of observed marked horizontal variation of the refractive index. The bending for these ray paths was then compared with values obtained under the normal assumption of horizontal homogeneity.

Although at 1 kilometer and above these horizontal changes appear to have little effect, rays emitted at low elevation angles are sensitive to extreme horizontal variations of the atmosphere near the surface, such as those associated with ducting conditions. However, since it appears that such conditions occur less than 15 percent of the time at most locations, the majority of ray-path calculations may be carried out under the normal assumption of horizontal stratification of the refractive index.

1. Introduction and Background

It is common in ray tracing studies to assume that the refractive index of the atmosphere is spherically stratified with respect to the surface of the earth [1, 2, 3, 4, 5].[2] Thus, the effect of refractive index changes in the horizontal direction is normally not considered, although recently Wong [6] has considered the effect of mathematically smooth horizontal changes in airborne propagation problems.

Neglecting the effect of horizontal gradients seems quite reasonable in the troposphere because of the relatively slow horizontal change of refractive index in contrast to the rapid decrease with height. In fact, examination of climatic data indicates that one must compare sea level stations located 500 km from each other on the earth's surface in order to observe a difference in refractive index values which would be comparable to that obtained by taking any one of these locations and comparing its surface value with the refractive index 1 km above the location. Although the assumption of small horizontal changes of the refractive index appears to be true in the average or climatic sense, there are many special cases such as frontal zones and land-sea breeze effects where one would expect the refractive index to change abruptly within the 80-odd kilometers of horizontal distance traversed by a tangential ray passing through the first kilometer in height.

It is these latter variations that are investigated in the present paper. Two cases of marked horizontal change of refractive index conditions were studied; one which occurred over the Canterbury Plain in New Zealand and the other at Cape Canaveral, Fla. Although these particular sites were chosen for several reasons such as land-to-sea paths and a subtropical location (where marked changes in refraction conditions are common) the major consideration was that detailed aircraft and ground meteorological observations were available for prolonged periods.

These detailed meteorological measurements allow a quantitative evaluation of the error apt to be incurred by assuming that the refractive index is horizontally stratified. The procedure used was to determine the refractive index structure vertically over the transmitter and assume that this same structure described the atmosphere everywhere. Rays were then traced through this horizontally laminated atmosphere. These ray paths were then compared with those obtained by the step-by-step ray tracing through the detailed convolutions of refractive index structure in the two cases under study.

In the sections that follow we will discuss the two cases chosen for study, the methods of calculations used to evaluate refraction effects, and the degree of confidence to which standard prediction methods may be used under conditions of horizontal inhomogeneity.

2. Canterbury

The Canterbury data was compiled by a radio meteorological team working from September 1946, through November 1947, on the South Island of New Zealand under the leadership of R. S. Unwin [7]. This report proved invaluable in this investigation as it was very carefully prepared, giving minute details of the experiment on a day-to-day basis.

[1] This work was sponsored in part by Task 31 of the U.S. Navy Weather Research Facility, Norfolk, Va.
[2] Figures in brackets indicate the literature references at the end of this paper.

Anson aircraft and a trawler were used for meteorological measurements over the sea, and three mobile sounding trucks for observations on land. The trucks and the trawler carried wired-sonde equipment, whereby elements for measuring temperature and humidity up to a height of 150 m to 600 m (depending on wind conditions) were elevated by means of balloons or kites. Standard meteorological instruments provided a continuous record of wind, surface pressure, temperature, and humidity at stations at the coast and 14 km and 38 km inland. The headquarters of the project were at Ashburton Aerodrome, and the observations extended out to sea on a line perpendicular to the coastline of Canterbury Plain. Aircraft were equipped with a wet- and dry-bulb psychrometer, mounted on the port-side above the wing. Readings were taken three or four times on each horizontal flight leg of 2 or 3 min duration. Special lag and airspeed corrections were applied, resulting in accuracy of $\pm 0.1°$ C. It was found that, under the variety of conditions in which observations were made, the aircraft flights were more or less parallel to the surface isobars; hence, the sea-level pressure as recorded at the beach site was considered to hold over the whole track covered by the aircraft. The relationship used for calculating the pressure; P, in millibars at a height h in feet was:

$$P(h) = P_0 - h/30$$

where P_0 is the surface pressure. This approximation (determined by averaging the effect of the temperature and humidity distributions on pressure in a column of air) resulted in a maximum error in the refractivity of 0.5 percent at 900 m. Radiosonde ascents at Hokitika on the west coast of South Island and Paraparaumuo and Auckland on North Island were used to supplement the aircraft measurements, particularly in the altitude levels above 1 km.

The observations, diagrams and meteorological records were studied, and a profile of unusually heterogeneous nature was chosen. The synoptic situation for the morning of November 5, 1947 was selected, as it revealed a surface-ducting gradient near the coast with an elevated layer about 100 km off shore. A cross section of the area from Ashburton to a point 200 km off shore was plotted with all available data, and isopleths of modified refractive index, M, were drawn to intervals of 2.5 units.

$$M = (n - 1 + K_e h)\ 10^6, \qquad (1)$$

where $K_e = (15.70)\ (10^{-8})$/m and

$$N = (n-1)10^6 = \frac{77.6}{T}\left[P + \frac{4{,}810 e_s\, RH}{T}\right], \qquad (2)$$

where P is the station atmospheric pressure in millibars, RH is the percent of the saturation vapor pressure, e_s, in millibars at the absolute temperature, T, in degrees Kelvin and n is the refractive index [8]. A simplified version of the lower portion of this cross section with the corresponding M curves is accompanied by a sketch of the general location of the experiment in figure 1. Some smoothing was necessary, particularly near the sea surface and in those areas where aircraft slant ascents and descents caused lag errors in altimeter readings and temperature and humidity elements. Isopleths over land were plotted above surface rather than above sea level with an additional adjustment in the scale ratios of height and distance in an attempt to simplify the reading of values from the diagram.

3. Cape Canaveral

The second area studied was the Cape Canaveral to Nassau path for the period of April 24 to May 8, 1957. This material was supplied by the Wave Propagation Branch of Naval Research Laboratories and the University of Florida. The particular case chosen for study was the meteorological profile of May 7, 1957, (2000 e.s.t.) due to its heterogeneous nature, showing a well-defined elevated layer

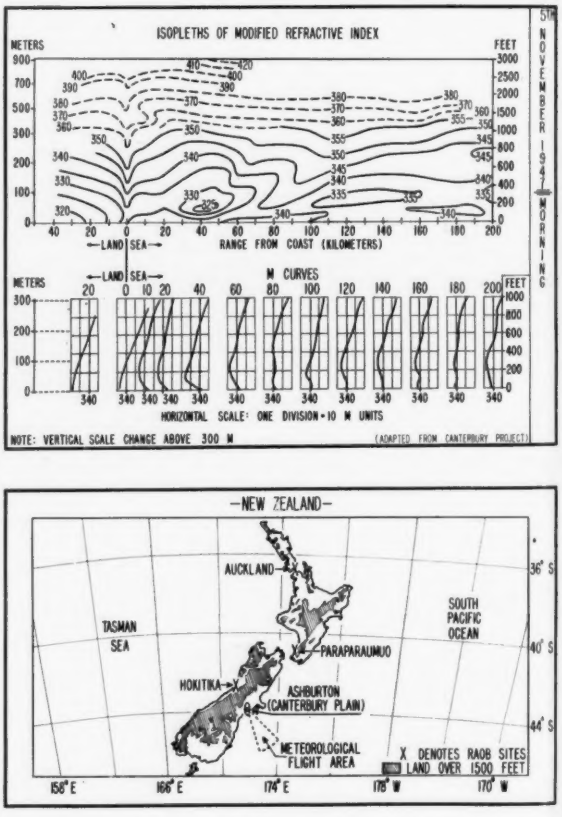


FIGURE 1. *Isopleths and curves of refractive index for Nov. 5, 1947, at Canterbury, with a map locating sources of meteorological data.* (*Courtesy of the Canterbury project.*)

at about 1,500 m. Fourteen refractometer soundings from aircraft measurements taken at various locations along the 487-km path (fig. 2) and six refractive index profiles (deduced from radiosonde ascents from Cape Canaveral, Grand Bahama Island, and Eleuthera Island) were read in order to plot a cross section of the atmosphere which would represent as closely as possible the actual refractive conditions at that time. Unfortunately, the data near the surface (up to 300 m) were quite sparse compared to those recorded in the Canterbury Project, and calibration and lag errors had not been noted as carefully in this preliminary report; therefore, some interpolation and considerable smoothing of refractive index values were necessary when drawing isopleths.

4. Ray Bending

The classic expression for the angular change, τ, or the bending of a ray passing from a point where the refractive index is n_1 to a second point where the refractive index is n_2 is given by [5]

$$\tau_{1,2}=-\int_{n_1}^{n_2}\frac{dn}{n}\cot\theta, \tag{3}$$

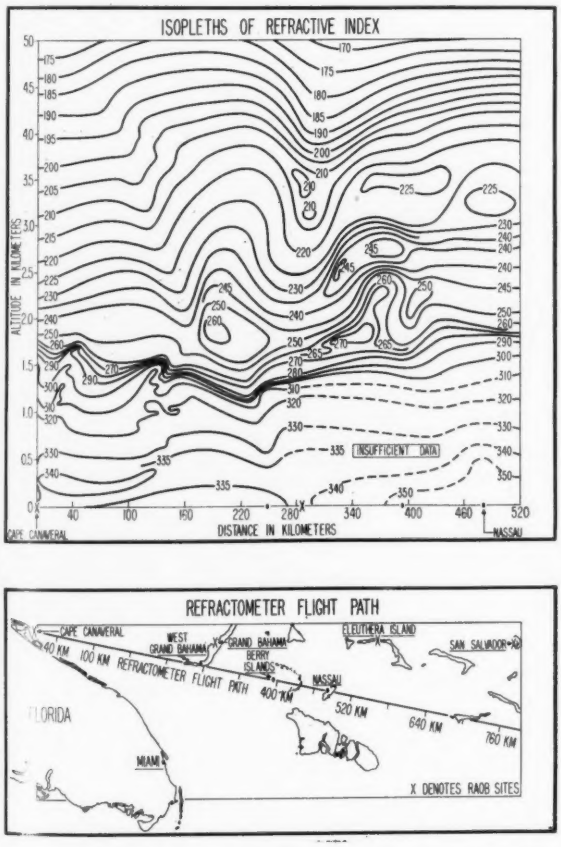


FIGURE 2. *Isopleths of refractive index and map of refractometer flight for May 7, 1957, Cape Canaveral to Nassau.*

where θ is the local elevation angle. Equation (3) was evaluated by use of

$$\Delta\tau_{1,2}=\frac{(N_1-N_2)}{\bar{n}_{1,2}}\times10^{-6}\cot\bar{\theta}, \tag{4}$$

where

$$\bar{\theta}=\frac{\theta_1+\theta_2}{2}.$$

The value of θ at each point was determined from Snell's law:

$$n_1\, r_1\cos\theta_1=n_2\, r_2\cos\theta_2=\text{constant}, \tag{5}$$

where r is the radial distance from the center of the earth and is given by $a+h$, where a represents the radius of the earth and h the altitude of the point under consideration. For simplicity one may rewrite (5) as

$$(1+N_1\times10^{-6})(a+h_1)\ \cos\ \theta_1=(1+N_2\times10^{-6})(a+h_2)\cos\theta_2. \tag{6}$$

Then, when θ is small, one may expand (6), neglect second order terms and obtain the convenient expression:

$$\theta_2=\left\{\theta_1^2+\frac{2(h_2-h_1)}{a}-2(N_1-N_2)\times10^{-6}\right\}^{1/2}, \tag{7}$$

where all values of θ are in milliradians.

After obtaining τ by use of (6) or (7), one may determine the distance, d, along the earth's surface that the ray has traveled from:

$$d_{1,2}=a[\tau_{1,2}+(\theta_2-\theta_1)]. \tag{8}$$

Thus by successive application of the above formulas, one may trace the progress of the radio wave as it traverses its curved path through the atmosphere. Normally the use of these equations is quite straightforward. When considering horizontal changes in n, however, one must satisfy these equations by iterative methods. In the present application, since n had to be determined by graphical methods, it was felt to be sufficient to assume a constant distance increment of 250 to 500 m, solve for appropriate height increment from

$$\Delta h=\Delta d\ tan\ \theta_1\left[1+\frac{h}{a}\right], \tag{9}$$

graphically determine N for the point $d_1+\Delta d$, $h_1+\Delta h$ and then determine θ_2 and $\tau_{1,2}$.

This latter type of ray tracing was done for various rays of initial elevation angles between 261.8 milliradians (mr) (15°) and 10 mr (~0.5°). The calculations were not carried to smaller elevation angles since this type of ray tracing is not valid within surface ducts for initial elevation angles below the angle of penetration [9, 10].

5. Comparisons

Although both of the calculated ray paths consisted of an oversea itinerary with coastal transmission sites, they are quite different in other aspects. Canterbury Plain is located southeast of the 10,000-ft chain of the Southern New Zealand Alps at a latitude of 44° S (the equatorward edge of the westerly belt of winds in November). Cape Canaveral is located on a sea level peninsula at 28° N (the poleward edge of the northeast trade circulation in May). While the Canterbury profile showed superrefractive tendencies, the Canaveral profile illustrated subrefraction at the surface counterbalanced by an elevated trade wind inversion layer, indicating that the total bending values of Canterbury would be higher than normal, while the Canaveral example would have values near or lower than normal.

These differences are illustrated by figures 3 and 4 where the bending, τ, in milliradians is plotted versus altitude in kilometers. The effect of horizontal changes is most pronounced for rays with initial elevation angles of 10 mr. On these figures the term "vertical" ray is used to designate the ray path through the horizontally homogeneous n structure determined from the refractive index vertically over the station. The term "horizontal" ray designates the ray path through the complex actual n structure. It is quite evident that a consistent difference in bending of about 1 mr exists between the "vertical" and "horizontal" rays at Canterbury above 1 km for $\theta_0=10$ mr. This would be expected since the vertical M profile (fig. 1) at the beach (our hypothetical transmitter site) is nearly normal in gradient while as little as 10 km to sea a duct exists, thus indicating a near maximum difference between the "horizontal" and the "vertical" rays at any initial elevation angle small enough to be affected by the duct. This is in contrast, however, with the case of Cape Canaveral where, except for the region of the elevated duct centered at about 1,500 m, the "vertical" and "horizontal" rays are in quite close agreement. These two examples illustrate that horizontal variations must be near the surface to be most effective. The importance of the altitude of the variation is due to the fact that refraction effects are very heavily weighted toward the initial layers [10].

Also shown on figures 3 and 4 are the values of the bending which would be predicted from the Central Radio Propagation Laboratory corrected exponential reference atmosphere [11]. The values shown are obtained from the value of N at the transmitter site as corrected by the vertical gradient over the first 100 m. It is noted that, for $\theta_0=10$ mr at Canterbury, the value of bending predicted by the model is in essential agreement with the "vertical ray" bending but underestimates the "horizontal ray" (which has the largest variation of n with horizontal distance) by about 1.25 mr. For Cape Canaveral at $\theta_0=10$ mr, the model atmosphere overestimates the bending by about 1.25 mr for altitudes in excess of 2 km. It should be emphasized that, although the model exponential atmosphere appears to represent the average of the two specific cases studied, the departure from this average arises from quite different causes in each case. The differences in the Canterbury case arise from the marked effect of horizontal variation of n as is indicated by the agreement of the vertical ray bending with the model atmosphere. The disagreement in the Canaveral case is due to the presence of a very shallow surface layer of nearly normal gradient topped by a strong subrefractive layer; therefore, it represents a shortcoming of the model rather than an effect of horizontal changes of n.

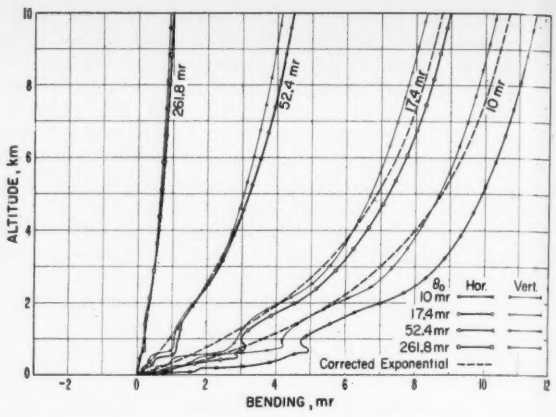


FIGURE 3. *Canterbury, 0 to 10 km, altitude versus ray bending.*

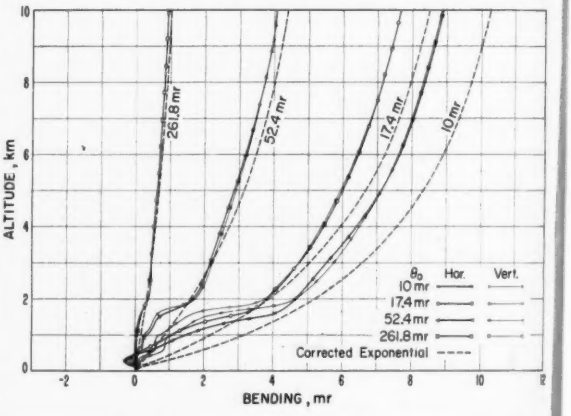


FIGURE 4. *Cape Canaveral, 0 to 10 km, altitude versus ray bending.*

The preceding analysis of bending throws the refractive differences in each case into sharp relief. The effect of refraction, of course, is to vary the ray path. Figures 5 and 6 show the ray paths corre-

sponding to the bendings of figures 3 and 4. Note that for Canterbury at $\theta_0=10$ mr the effect of the horizontal variation of n is to produce a difference in estimation of about 1 km in height or 20 km in ground distance at 300 km from what would be obtained from considering the vertical n profile as a representation of the entire path. The effect of the subrefractive layer at Cape Canaveral is not so large, but it does cause an overestimation of the ground distance by about 5 km and an underestimation of the height by less than one quarter of a kilometer at a ground distance of 300 km by assuming that the vertical profile may be used throughout the entire ray path.

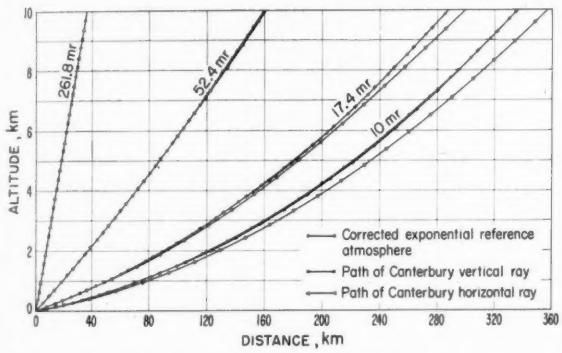


FIGURE 5. *Canterbury, 0 to 10 km, altitude of ray versus distance.*

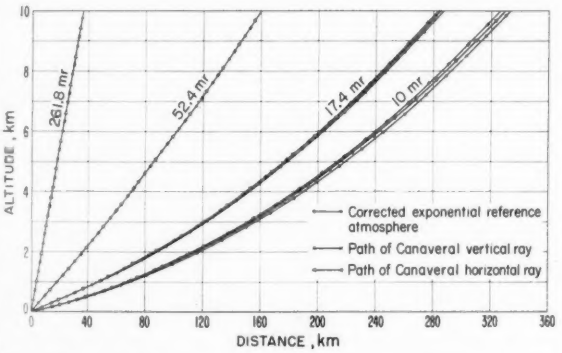


FIGURE 6. *Cape Canaveral, 0 to 10 km, altitude of ray versus distance.*

6. Extension to Other Regions

It should be pointed out that the ducting case at Canterbury represents an extreme refraction condition and is not necessarily typical of conditions observed in other regions or, indeed, at Canterbury. The Canterbury project was purposely restricted to a study of ducting conditions with the result that less than 20 percent of the total observations for the fifteen months are reported. Therefore, because one of the more extreme cases is represented by the November 5th example, one might conclude that much less than 20 percent of the observations would show the same degree of horizontal n change as the profile studied.

If one further hypothesizes that the greatest horizontal n change would be associated with ducting conditions, then the percentage incidence of ducts as evaluated from radiosonde observations, listed for various stations in table 1, would indicate that the effects of horizontal changes of n sufficient to cause variations in the ray path as large as those of the present study would be observed less than 15 percent of the time, regardless of geographic location.

TABLE 1. *Percentage occurrence of surface ducts during the years 1952 to 1956*

Station	% incidence			
	February	*May*	*August*	*November*
Fairbanks, Alaska	9.4	0.4	0.4	6.2
Columbia, Mo.	.7	2.5	8.4	1.3
Washington, D.C.	.7	4.8	4.3	1.4
Canton Island	10.0	9.2	12.4	11.5
Miami, Fla.	0.7	3.5	8.5	2.7

The probable importance of subrefractive layers upon the prediction of refraction effects has emerged as a secondary result of the present study. Although subrefraction is normally neglected, it is potentially a very important refractive factor for distances of, say, less than 40 km. Even though the percentage occurrence of subrefractive layers can be as large as 6 percent (see table 2), this effect is frequently offset by the concurrent occurrence of an adjacent superrefractive layer, as is illustrated by the Cape Canaveral example.

TABLE 2. *Percentage occurrence of surface subrefractive layers during the years 1952 to 1956*

Station	% incidence			
	February	*May*	*August*	*November*
Fairbanks, Alaska	0.0	0.0	1.2	0.4
Columbia, Mo.	.0	1.6	0.6	4.0
Washington, D.C.	.9	2.2	5.8	2.7
Canton Island	.0	0.0	0.0	0.3
Miami, Fla.	.7	.3	.9	.7

7. Conclusions

The conclusions of the present study could be considerably modified by the analysis of many more examples, although it is evident that horizontal variation of n near the earth's surface produces the most marked deviations from the ray paths obtained by assuming horizontal stratification of n. The effect of horizontal changes occurring more than a kilometer above the surface appear from our present examples, to have little effect. Further, the effects

of horizontal changes appear to be most pronounced in the presence of surface ducts and at small elevation angles. The tentative conclusion is reached that the effect of horizontal n change is normally small since ducting will occur less than 15 percent of the time.

The authors express their appreciation to C. M. Miller, G. E. Richmond, B. J. Weddle, and P. C. Whittaker for their aid in the laborious calculations necessary for this study.

In addition, the authors express their gratitude for permission to use the data of the Canterbury Project and the Wave Propagation Branch of the Naval Research Laboratories. Indeed, the present study would have been quite impossible without their unusually detailed reports.

8. References

[1] L. J. Anderson, Tropospheric bending of radio waves, Trans. Am. Geophys. Union **39,** 208 (1958).

[2] G. H. Millman, Atmospheric effects on VHF and UHF propagation, Proc. I.R.E. **46,** 1492 (1958).

[3] B. M. Fannin and K. H. Jehn, A study of radar elevation angle errors due to atmospheric refraction, I.R.E. Trans. **AP-5,** 71 (1957).

[4] B. R. Bean and B. A. Cahoon, The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles, Proc. I.R.E. **45,** 1545 (1957).

[5] M. Schulkin, Average radio-ray refraction in the lower atmosphere, Proc. I.R.E. **40,** 554 (1952).

[6] M. S. Wong, Refraction anomalies in airborne propagation, Proc. I.R.E. **46,** 1628 (1958).

[7] Report of factual data from the Canterbury project, vols. I–III, Dept. Sci. Ind. Research, Wellington, New Zealand (1951).

[8] E. K. Smith and S. Weintraub, The constants in the equation for atmospheric refractive index at radio frequencies, Proc. I.R.E. **41,** 1035 (1953)

[9] D. E. Kerr, Propagation of Short Radio Waves, Radiation Lab. Ser. **13,** 17 (McGraw-Hill Book Co., New York, N.Y., 1951).

[10] B. R. Bean, Climatology of ground-based radio ducts, J. Research NBS **63D,** 29 (1959).

[11] B. R. Bean and G. D. Thayer, On models of the atmospheric radio refractive index, Proc. I.R.E. **47,** 740 (1959).

Boulder, Colo. Paper 63D3–24

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

On the Correlation of Solar Noise Fluctuations in Harmonically Related Bands

L. R. O. Storey

(May 25, 1959)

A method is proposed for the study of the solar corona, by observing a delayed correlation between rapid fluctuations of enhanced solar radio emission in harmonically related frequency bands. The correlation is expected in those types of emission that are produced by nonlinear plasma oscillations. The delay of the fundamental frequency with respect to the harmonic would be brought about by dispersive group retardation in the corona. The method appears to be most suitable for use with type II bursts, though it might also be applied to other types of nonthermal solar emission.

A solar noise burst of type II consists of a narrow band of intense radio emission, that drifts steadily downwards in frequency over a period of several minutes (Wild and Macready [1] and[1] Wild [2]). Often two such bands are observed simultaneously, with center frequencies roughly in the ratio 1:2. This fact suggests that the burst originates in nonlinear plasma oscillations in the solar corona (Wild, Murray, and Rowe [3]; Maxwell, Swarup, and Thompson [4]). The oscillations are thought to be excited by the motion through the corona of a bunch of particles that has been ejected violently from the photosphere. Usually the burst is preceded by a solar flare which, in these cases, may well be the original source of the fast particles. At each point along their course outward through the corona, the impact of the particles excites oscillations at the local plasma frequency, which radiate the noise observed in the lower of the two bands. Because the oscillations are supposedly nonlinear, energy is radiated also at the second harmonic of the plasma frequency, and this corresponds to the upper band. As the bunch of particles continues outward, it encounters progressively decreasing local plasma frequencies, and this circumstance explains the downward progression of the frequency of the burst.

Now it is possible that the noise in these bursts may exhibit fairly rapid nonstatistical fluctuations, of duration 0.1 sec or less, due to irregularities either in the distribution of electrons in the corona or in the strength of the source of excitation for the oscillations. If so, then the same fluctuations would be expected to occur on both the fundamental frequency and on the harmonic, since the two frequencies are generated simultaneously at the same point by the same process. However, the fluctuations would not be exactly simultaneous when observed at the earth, because, as the waves travel outward through the corona, the different frequencies would experience different amounts of group delay; the fundamental would be delayed more than the harmonic. If the differential delay between the fluctuations on the two frequencies could be observed and measured, it might prove to be a source of useful information about the solar corona.

To find out whether this differential delay is likely to be observable, its size has to be estimated. In the appendix, some calculations are made on the assumptions that the electrons in the corona are distributed with spherical symmetry about the center of the sun, and that their number density N varies with distance from the center according to the power law

$$N=N_0R^{-b}, \tag{1}$$

where N_0 is the nominal density at the surface of the photosphere, and the distance R is measured in solar radii. The source of excitation for the burst is assumed to be traveling radially outwards from the sun, along the line directly towards the earth. In calculating the group delays, no account was taken of the possibility that magnetic fields might be present. On these assumptions, the differential delay between a fluctuation on the fundamental frequency f and the same fluctuation on the second harmonic ($2f$) is

$$\Delta T(f)=Kf^{-2/b}, \tag{2}$$

where the factor of proportionality K depends on the values of N_0 and b. These parameters were estimated by fitting the distribution (1) to the densities given by Van de Hulst [5], which were obtained experimentally by measurement of the brightness and polarization of coronal light at different angular distances from the sun. The densities used were those that apply at sunspot maximum, and the fit was made at the level where the density corresponds to a plasma frequency of about 100 Mc; this process gave the values $N_0=4\times10^8$ cm^{-3}, and $b=9$. Thence $K\approx1$, when ΔT is measured in

[1] Figures in brackets indicate the literature references at the end of this paper.

seconds, and f in megacycles. Equation 2 then yields the following estimates of the delay at various frequencies of the fundamental:

f	ΔT
80 Mc	0.40 sec
100 Mc	.38 sec
120 Mc	.36 sec

These are lower estimates because, above the 100 Mc level, the densities given by (1) are rather less than those of Van de Hulst's distribution; yet, if that distribution is correct, then they are not likely to be excessively low, as most of the differential delay of the waves is produced just above their level of origin.[2]

It appears that a differential delay of about 0.4 sec should be expected, and that, for the assumed distribution, it should vary only slowly with frequency. A delay of this order should be observable readily, provided that the noise itself contains fluctuations of much shorter duration. If observed on just one harmonically related pair of frequencies, the delay would give a weighted measure of the total electron content of the corona above the level of origin of the waves. If the delay could be measured at several such pairs of frequencies, and the same assumptions were made as in the calculations above, then the observed variation with frequency would give the value of the index b in the power law for the distribution of electrons. More generally, it should be possible to infer the distribution of electrons in height with no prior assumptions about its form, by a variant of the "lamination" methods that have been developed for analyzing the group delays observed in ionospheric sounding (Thomas [8]). This information, when combined with the observed frequency/time relation for the burst, would show how the height of the source of excitation in the corona increased with time. At present, the radial movement of the source can be traced only by assuming some model for the distribution.

In this kind of analysis, height would be measured with respect to the level of origin of the highest fundamental frequency that was observed. The height of this particular level with respect to the photosphere would have to be estimated in some other way. For instance, if the radio burst was preceded by a flare, then the time interval might be measured between the occurrence of the flare and the onset of the burst; this information, when combined with the calculated velocity of movement of the source of excitation, would give the height.

Such an analysis would also be complicated by the fact that the source would not, in general, be moving directly towards the earth; nor would it necessarily, even, be moving radially outwards from the sun (Wild, Murray, and Rowe [3]). However, it is worth noting that any motion of the source transverse to the line of sight can be observed directly by other means (Payne-Scott and Little [7]). If such observations were combined with those of the differential delay, it might then be possible to infer the distribution on the assumption of spherical symmetry. This assumption itself may seem naïve, considering that the corona is known to have an irregular structure, but nonetheless it may still be useful as a basis for deriving a first approximation to the distribution.

The first step in implementing these proposals would be to find out whether the noise in the type II bursts does exhibit nonstatistical fluctuations that are sufficiently rapid to be of use for measuring the differential delay. This question cannot be answered from the results of previous observations, since all these were made using receivers with too large bandwidths. In detecting a signal that varies in frequency, the best bandwidth to use is roughly equal to the square root of the rate of change of frequency (Storey and Grierson [10]). Judging from the spectral characteristics of the type II bursts, the best frequency range in which to observe them on the fundamental would be 80 to 120 Mc, and in this range their rate of change of frequency is usually about 1 Mc[2]. Hence the RF bandwidth of the receiver should be about 1 kc. With so narrow a bandwidth, a fairly large antenna would be required to provide adequate sensitivity.

If rapid fluctuations were observed, the next step would be to search for the anticipated correlation between the fluctuations in harmonically related bands. The antenna would be connected to a pair of receivers, one tuned to a fixed frequency in band 80 to 120 Mc, and the other to twice this frequency; the high-frequency receiver should have an RF bandwidth double that of the low-frequency receiver (the criterion of rate of change of frequency would require the bandwidth at the harmonic to be $\sqrt{2}$ times that at the fundamental, but this requirement is not at all critical, so it seems better to arrange that the limits of the two bands are related harmonically). The outputs from the detectors of the two receivers would be recorded on magnetic tape, together with timing signals. If a particular section of tape was identified as the record of a type II burst, then this record would be played back at reduced speed, the information on the two data tracks converted to digital form, and the cross-correlogram of the two sets of fluctuations evaluated in a digital computer. In this way a search would be made for a correlation between the fluctuations observed in the two bands and for a possible systematic time displacement that would give the best correlation. If this search were to prove successful, then the measurements would be extended to cover more than one pair of frequencies. There would be advantage to making these measurements in conjunction with recordings on a panoramic radio spectrograph, which could be used to identify the type II bursts.

[2] However, some recent studies indicate that Van de Hulst's figures may be too low. Newkirk [6], on the basis of further optical measurements, concludes that even on the quiet disk the figures should be multiplied by a factor of 2.2, while over active regions the factor should be 4.5. Maxwell [7], on evidence from the radio bursts themselves, has suggested that the multiplying factor for active regions may be as high as nine. If such factors indeed apply, then the estimates of the differential delay ΔT should be multiplied by these same factors raised to the power of $1/b$ (see eq 11 in the appendix); for the given value of b, the increase in the estimates is less than 25 percent.

Possibly the same techniques could be used to study other types of enhanced solar radiation. Thus the noise bursts of type III also exhibit harmonic structure (Wild, Murray, and Rowe [3]), but the rate of change of frequency is much faster, of the order of 50 Mc²; hence the receiver would need to have a rather larger RF bandwidth, about 7 kc, for recording these bursts. A further possibility is the study of noise storms. Though these disturbances are not localized in frequency at any one time, as are the bursts, but to the contrary are spread over a wide frequency band (Wild and Macready [1]), nevertheless they also may be generated by the same process involving nonlinear plasma oscillations. The difference between the two kinds of disturbance may lie only in the nature of the excitation; a continuous stream of particles for the storm, as against a discrete bunch for the burst. If so, then there would be the same grounds for expecting a delayed correlation between fluctuations of noise power in harmonically related bands; in this case, however, there is no way of telling in advance what RF bandwidth would be best for observing the correlation. I am indebted to R. S. Lawrence for this suggestion.

In this work, there is one possible source of difficulty that needs to be recognized. It is the presence of irregularities in the corona, which may cause the radiation to be propagated from the source to the earth by more than one path. Multipath propagation could give rise to rapid and severe fading that would not be correlated in harmonically related frequency bands, and any correlated fluctuations might thus be masked. A similar difficulty has been met with already in an attempt to measure the electron density at great heights above the earth by observing the dispersion of radar echoes from the moon (Eshleman, Gallagher, and Barthle [11]). However, it is hard to judge whether this factor will present a serious obstacle to the proposed studies.

In writing this paper, I have been helped considerably by advice from Dr. Constance S. Warwick of this Laboratory, and from Dr. G. Newkirk of the High Altitude Observatory of the University of Colorado.

Appendix: Calculation of the Differential Delay

In calculating the differential delay, the following symbols will be used to refer to points, first, at the surface of the photosphere, and second, at some level in the corona:

	Photosphere	*Corona*
Distance from center of sun	r_0	r
Local electron density	N_0	N
Local plasma frequency	F_0	f_p

The electron density in the corona will be assumed to depend on r only, according to the power law

$$N=N_0R^{-b} \tag{1}$$

where

$$R=r/r_0 \tag{3}$$

The plasma frequency f_p is proportional to the square root of N, so its distribution is

$$f_p=f_0R^{-b/2}. \tag{4}$$

By hypothesis, the oscillations that produce the noise emissions take place at the local plasma frequency; emissions of frequency f are produced at the level, $r(f)$ say, where $f_p=f$. The difference in the group delay experienced by the fundamental of frequency f and the harmonic of frequency $2f$, in traveling from this level out through the corona to the earth, is

$$\Delta T(f)=1/c\int_{r(f)}^{\infty}\{\mu'(f)-\mu'(2f)\}dr \tag{5}$$

where c is the speed of light, and μ' is the group refractive index. If magnetic fields are absent, as will be assumed, then μ' depends only on the relative values of the wave frequency and local plasma frequency, according to the law

$$\mu'(f)=(1-X)^{-\frac{1}{2}} \tag{6}$$

where

$$X=(f_p/f)^2 \tag{7}$$

Along the path of the waves, the dimensionless quantity X varies from a value of unity at the source to zero at infinity. It is convenient to adopt X as the variable of integration:

$$\Delta T(f)=1/c\int_1^0\{\mu'(f)-\mu'(2f)\}\left(\frac{dr}{dX}\right)dX$$
$$=1/c\int_0^1\{\mu'(f)-\mu'(2f)\}\left(-\frac{dr}{dX}\right)dX. \tag{8}$$

The components of the integrand are obtained as follows: first, from (6),

$$\mu'(f)-\mu'(2f)=(1-X)^{-\frac{1}{2}}-(1-X/4)^{-\frac{1}{2}}. \tag{9}$$

Also, from (7), (4), and (3),

$$-\frac{dr}{dX}=r_0/b\,(f_0/f)^{2/b}X^{-(1+1/b)}. \tag{10}$$

Hence the differential delay is

$$\Delta T(f)=Kf^{-2/b} \tag{2}$$

where

$$K=\frac{r_0f_0^{2/b}}{bc}I \tag{11}$$

and

$$I=\int_0^1\{(1-X)^{-\frac{1}{2}}-(1-X/4)^{-\frac{1}{2}}\}X^{-(1+1/b)}dX. \tag{12}$$

Now the factor K must be evaluated. The constants are $c=3\times10^5$ kmps, and $r_0=7\times10^5$ km, while, from the data of Van de Hulst [5], $b\simeq9$ and $f_0\simeq180$ Mc. The integral I can be evaluated numerically, though with some difficulty, since one or other of the factors in the integrand becomes infinite at each limit. The result is that, when f is measured in megacycles, $K=1$. (1).

References

[1] J. P. Wild and L. L. Macready, Observations of the spectrum of high-intensity solar radiation at metre wavelengths, Part I—The apparatus and spectral types of solar burst observed, Australian J. Sci. Research [A] **3**, 387 (1950).

[2] J. P. Wild, Observations of the spectrum of high-intensity solar radiation at metre wavelengths, Part 2—Outbursts (type 2), Australian J. Sci. Research [A] **3,** 399 (1950).

[3] J. P. Wild, J. D. Murray, and W. C. Rowe, Harmonics in the spectra of solar radio disturbances, Australian J. Phys. **7,** 439 (1954).

[4] A. Maxwell, G. Swarup, and A. R. Thompson, The radio spectrum of solar activity, Proc. Inst. Rad. Engrs. **46,** 142 (1958).

[5] H. C. Van de Hulst, *in* The Sun (ed. G. H. Kuiper) p. 262, table 10 (Chicago, 1953).

[6] G. Newkirk, A model of the electron corona with reference to radio observations, High Altitude Observatory, U. of Colorado, Solar Research Memo No. 109.

[7] A. Maxwell (private communication).

[8] J. O. Thomas, The distribution of electrons in the ionosphere, Proc. Inst. Rad. Engrs. **47,** 162 (1959).

[9] R. Payne-Scott and A. G. Little, The position and measurement on the solar disc of sources of radiation at a frequency of 97 Mc. Part 3—Outbursts, Australian J. Sci. Research [A] **5,** 32 (1952).

[10] L. R. O. Storey and J. K. Grierson, Time-symmetric filters, Part 2—Response to gliding tones, Electronic Eng. **30,** 648 (1958).

[11] V. R. Eshleman, P. B. Gallagher, and R. C. Barthle, On the use of HF moon echoes to determine cislunar ion densities, presented at URSI Spring Meeting (May 5, 1959).

BOULDER, COLO.

(Paper 63D3–25)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

A Monochromatic Low-Latitude Aurora

F. E. Roach and E. Marovich

(May 28, 1959)

A monochromatic (6300 A) auroral arc occurred over Colorado on September 29/30, 1957. It seems to have been a continuation of a similar arc observed at Haute Provence, France, on the same night. Its intensity decreased during the night from about 7,000 rayleighs to 2,000 rayleighs compared with a normal zenith intensity of 100 to 200 rayleighs. It was relatively fixed (geographically) during the night, south of and apparently independent of a visual aurora that was active to the north from 0100 to 0400 m.s.t. The magnetic dipole lines of force from the arc extend out into space between the two Van Allen radiation belts. It is speculated that the arc may be associated with one of the belts.

1. Introduction

In a recent study of low-latitude auroral activity, Barbier [1][1] has reported on observations at the Haute Provence Observatory (longitude 5°43′ E; latitude +43°56′; geomagnetic latitude 45.9°). He found evidence for a "monochromatic" (6300 A) aurora during 31 nights over the period from March 10, 1956 to September 29, 1957. In general, these auroras were not conspicuous visually because of the relative insensitivity of the eye to the red.

Barbier operates three different photometers at Haute Provence. One utilizes eight different filters and is fixed in orientation toward the celestial pole (with reference observations at the zenith). The other two cover the sky in almucantars, usually at a zenith distance of 75°. The almucantar sweeps show two distinct types of 6300 auroral activity: (1) a pronounced increase in brightness in the general northerly direction; and (2) an increase in brightness at two different azimuths—one a little north of west and the other a little north of east. The second type may be accompanied by an increase in the brightness at the celestial pole but not appreciably at the zenith. Barbier's interpretation of these double maximums is that there is a 6300 arc extending roughly from east to west across the sky, crossing the meridian somewhere between the zenith and the northern horizon.

On one of the four nights during which such an arc was observed at the Haute Provence Observatory (September 29/30, 1957), we obtained independent evidence at Fritz Peak (longitude 105°29′25″ W; latitude 39°54′34″ N; geomagnetic latitude 49°) that a similar 6300 arc occurred. We were observing 6300 along the geographic meridian from horizon to horizon and were not operating a 6300 photometer in almucantars. We did have 5577 photometers observing the sky in a series of almucantars both at Fritz Peak and at the Rapid City ION Y station (longitude 103°18′50″ W; latitude 44°01′ N; geomagnetic latitude 53°).

[1] Figures in brackets indicate the literature references at the end of this paper.

During the night a visual aurora was observed to the north of both American stations with a peak intensity at approximately 0215 m.s.t. A striking feature of the night was the isolation of a very pronounced maximum of intensity of 6300 near the zenith at Fritz Peak, as illustrated in figure 1. At the start of the night, the observer (L. R. Megill) noted that the 6300 records were off scale near the zenith and after moonset (2200 m.s.t.) he was convinced that an unusual event was occurring. The sensitivity of the photometer was reduced to keep the peak zenith deflection on scale. Visual inspec-

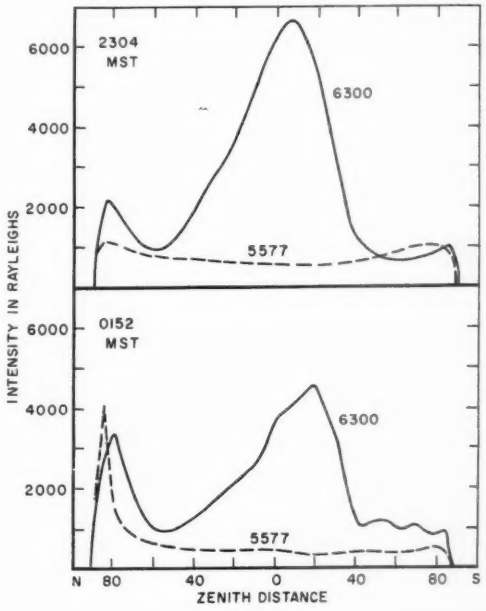


FIGURE 1. *Meridian sweeps at Fritz Peak of 6300 and 5577 at 2304 m.s.t. and 0152 m.s.t.*

tion of the sky did not show any apparent zenith brightening, which subsequently was explained as due to the fact that, in agreement with Barbier's interpretation, this was a monochromatic emission (5577 for example showed no zenith enhancement at all) and, though physically bright (about 7,000 rayleighs), was nevertheless below the visual threshold for red light.[2] The usual zenith intensity of 6300 is 100 to 200 rayleighs so that the observed arc is some 50 times normal.

From our observations alone, it is not possible to infer an arc structure from the zenith enhancement of 6300 in horizon to horizon sweeps in the meridian. But it is reasonable to assume such a structure when our records are interpreted in the light of Barbier's results.

2. Variation of the Peak Intensity with Time

The 6300 arc at Fritz Peak slowly decreased in intensity during the night. In figure 2, we show the variation of the peak intensity of 6300 above the background and, for comparison, the results of Barbier at Haute Provence. We note that the end of the observing night at Haute Provence coincides almost exactly with the start at Fritz Peak.

The physical picture suggested by these observations is that a 6300 arc persisted in time over at least 12 hr and in space at least 110° in longitude or approximately one-third of the way around the world (fig. 3).

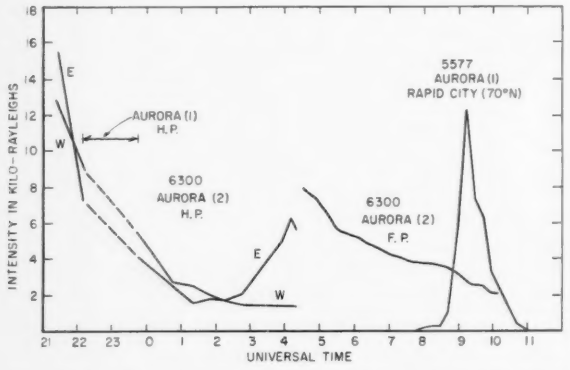


FIGURE 2. *Variation of intensity of the 6300 aurora at Haute Provence and Fritz Peak on September 29/30, 1957.*

A localized northern aurora at Haute Provence is indicated between 22h and 24h u.t. (Sept. 29) and at Rapid City, north of Fritz Peak, between 8h and 11h u.t. (Sept. 30).

3. Position of the Maximum

We have measured the zenith distance of the maximum intensity on our records with the results shown in table 1 (col. 2) and in figure 4. There is a

[2] It has been our experience that 1,000 rayleighs is the effective visual threshold for the green part of the spectrum. At 6300 A the threshold should be much higher than in the green since the dark-adapted eye is almost inert to the longer wavelengths.

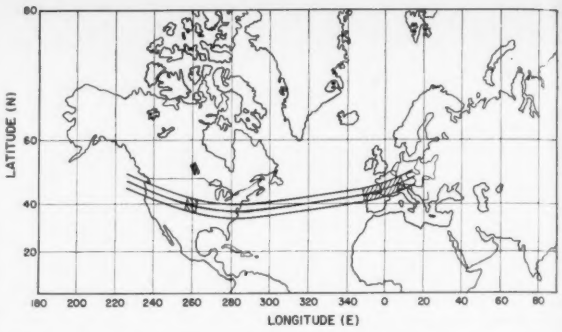


FIGURE 3. *Suggested extent of 6300 arc on September 29/30, 1957.*

Shaded portions are actually observed at Haute Provence and Fritz Peak.

TABLE 1

Mountain standard time	Position of maximum—degrees south of zenith	Intensity of maximum	Mountain standard time	Position of maximum—degrees south of zenith	Intensity of maximum
2134	...	(7870)	0030	5.0	4000
2142	...	(7630)	0038	4.0	3790
2150	...	(7460)	0046	4.5	3770
2158	...	(7410)			
2206	...	(6850)	0054	4.5	3800
			0102	5.5	3710
2214	...	(6390)	0110	6.0	3760
2222	...	(6230)	0118	7.0	3680
2230	3.5	5430	0126	8.0	3710
2238	3.0	5390			
2246	4.0	5390	0134	10.0	3710
			0142	8.0	3550
2254	5.5	5420	0150	14.0	3490
2302	6.0	5290	0158	13.0	3160
2310	7.0	5120	0206	11.5	2980
2318	6.5	5000			
2326	5.5	4810	0214	9.5	2800
			0222	16.0	2420
2334	7.5	4690	0230	14.5	2510
2342	5.5	4520	0238	14.0	2650
2350	5.5	4360	0246	10.5	2390
2358	5.0	4300			
0006	1.0	4260	0254	15.0	2140
			0302	13.0	2030
0014	3.0	4140	0310	16.0	2020
0022	4.0	4040			

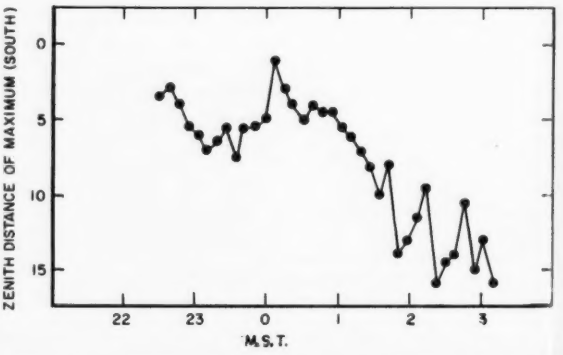


FIGURE 4. *Variation of position of the 6300 peak observed at Fritz Peak during the night of September 29/30, 1957.*

systematic southerly movement from midnight to 03^h. If this is interpreted as a geographical movement and if we assume an emission height of 300 km, the total geographical excursion of the photometric center of the arc is only about 80 km, or less than 1° of latitude (corresponding to a change of zenith distance from 1° S to 16° S). Between midnight and 03^h, the mean speed of movement was about 6.5 m/sec and again assuming a 300-km height.

The center of the arc in Colorado was within 1° of the geomagnetic latitude of Fritz Peak (49°) during the entire night. If the center of the arc at Haute Provence was at a northerly zenith distance of about 50°, it would be about 3° north of Haute Provence (Φ=46°) or also at a geomagnetic latitude of 49°.

The southern migration of the peak may be the result of the fact that the northern side of the photometric curve steadily moves southward. The southern side seems to remain constant. The result is an apparent movement of the peak (fig. 5).

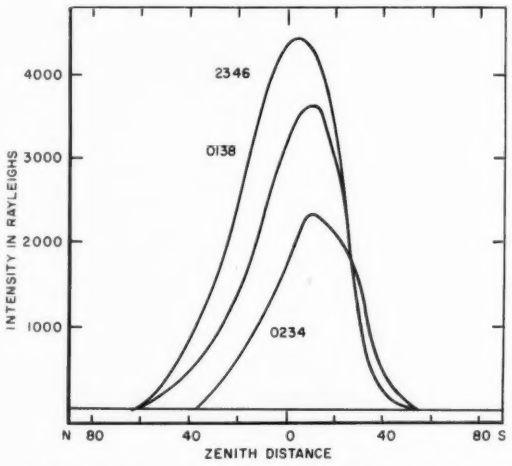


FIGURE 5. *Evolution of the 6300 arc at Fritz Peak during the night of September 29/30, 1957.*

4. Width of the Arc

The southern and northern extremes of the arc are fairly well defined on the records; hence, the extreme width of the arc can be estimated. If the height is fixed at 300 km, the mean width is 800 km. The half-intensity width is about 250 km. Barbier estimated the arc in France to have an extension of about 500 km. In section 7 we suggest however that there is a significant extension in height in which case these width estimates are too high.

5 Is the 6300 Arc Associated With the Northerly Aurora?

There was a well-defined aurora (which we shall designate Aurora (2)) near the northern horizon at both Haute Provence and Fritz Peak during the night in question as indicated in figure 1. So far as the intensity of the 6300 arc is concerned, there is no evidence at either station that it is influenced by the coincident northern aurora. The northern (5577) aurora peaks beyond the extreme (z=80°) observation at Rapid City which places it at least 470 km or 4.2° to the north. This corresponds to Φ=57°. The two phenomena are thus separated by at least 8° of latitude (≈900 km.)

Although the center of Aurora (2) in America is well to the north of even the northern extreme of the red arc, it is possibly significant that the southern edge of Aurora (2) is very close to the northern edge of the red arc (see fig. 6). The occurrence of the two phenomena on the same night and with their edges almost touching (geographically) suggests that there may be some connection between them. However, there is no definite evidence favoring any relationship and, as Barbier has indicated, "the evolution of the arc is not at all influenced by the presence of another concomitant auroral phenomenon."

FIGURE 6. *Relationship of the 6300 arc to the visual aurora.*

The shaded region is the 6300 arc and the ⊥ ⊥ symbols indicate the southern extension of the visual aurora.

That the 6300 monochromatic auroras at Haute Provence are positively correlated with geomagnetic activity is shown by Barbier. However, he calls attention to the fact that at Tamanrasset (longitude 5° 32′ E; latitude +22° 47′; geomagnetic latitude +25.3°) there are many irregularities in the 6300 intensity. In the course of 2 or 3 hr, an increase by as much as a factor of 10 may occur. These are similar to, but much larger than, the variations which accompany the auroras at Haute Provence, but they occur more frequently and have no relationship to magnetic activity.

According to Chapman [2] who has made a study of the occurrence of *tropical* auroras, their color is often conspicuously red. On one occasion (February 14, 1848) an aurora was observed in India at a geomagnetic latitude of $+13°$ at a time when there was no outstanding magnetic activity and no world-wide auroral display. We raise the question whether the tropical auroras (India, Tamanrasset) and the detached 6300 arcs, as observed at Haute Provence and Fritz Peak on September 29/30, 1957, may be similar phenomena.

6. A Possible Double-Aurora Mechanism

The narrowness of the 6300 arc suggests an analogy with localized features of the aurora. It is therefore of interest to consider it in relationship to the earth's magnetic field. In figure 7 we show an extension of the magnetic dipole lines of force for $\Phi=48.5°\pm3.0°$ from the earth's surface out into space (region B). The two regions of concentrated radiation reported by Van Allen, McIlwain, and Ludwig [3] (A and C) are also shown, along with the magnetic line of force corresponding to the auroral zone ($\Phi=67°$). It is noted that the region B bears the same relationship to A as D to C. This suggests the possibility that the inner region of ionization (A) may be responsible for our 6300 arc and the outer region (C) for the more northerly aurora. If region C received an energy surge simultaneously with A then the observations indicate that the release in the auroral zone was more sharply dependent on local time than that producing the 6300 arc.

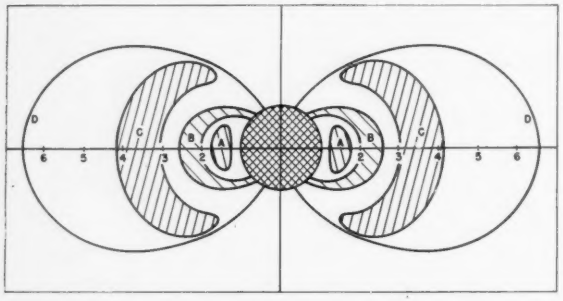


FIGURE 7. *The extension of the magnetic dipole field lines (B) corresponding to the terrestrial position of the 6300 arc.*

(A) and (C) are the Van Allen regions of enhanced ionization. *D* is the extension of the dipole line of force for the auroral zone ($\Phi=67°$).

The rate of decrease of intensity, I, noted by Barbier at Haute Provence is about twice that observed at Fritz Peak. Defining α by

$$\alpha=\frac{-d\log_e I}{dt},$$

he found $\alpha=1.1\times10^{-4}$ sec^{-1}. We find for Fritz Peak $\alpha\approx0.6\times10^{-4}$ sec^{-1}. The rate of decrease of 6300 during the night under nonauroral conditions is comparable [4] with $\alpha=0.4\times10^{-4}$ sec^{-1}.

7. Photometric Asymmetry of the Arc

A characteristic of the arc as observed at Fritz Peak is that the southern edge is steeper than the northern. This is the sort of effect that could be expected if, in addition to an extension in latitude, there is a significant extension along the magnetic lines of force as shown in figure 8. In both of the models shown in figure 8, the maximum contribution to the brightness is assumed to occur at a height of 300 km. In (a) the arc is shown with an extension in both altitude, (from 120 to 500 km) along the geomagnetic lines of force, and latitude (240 km). In (b) a thin layer (50 km) is assumed with a horizontal extension of 560 km.

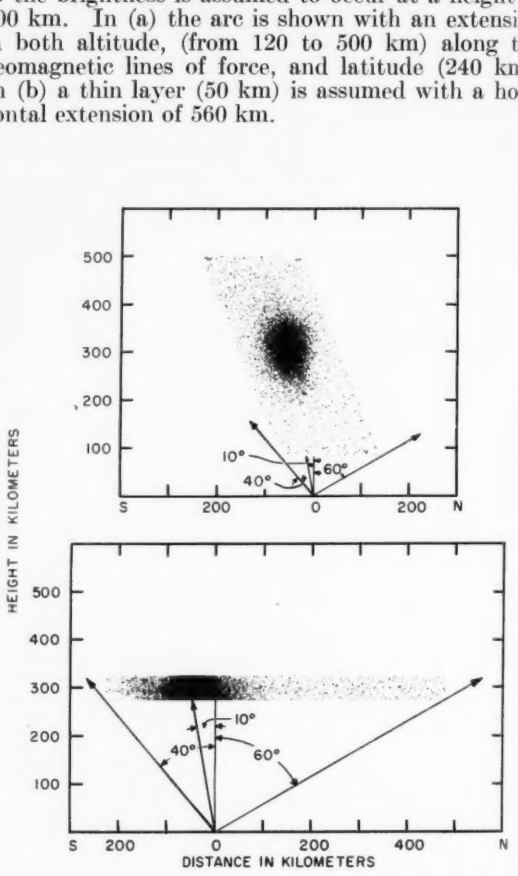


FIGURE 8. *Two schematic models of the 6300 arc based on (a) peak intensity 10° S, (b) southern extreme at 40° S, and (c) northern extreme at 60° N.*

8. Conclusion

We have made a preliminary examination of our records for similar monochromatic auroras on other occasions. On at least three other nights (September 21/22, September 22/23, and September 30/October 1, all in 1957), we have noted a considerable 6300 activity at Fritz Peak. There is evidence that the 6300 emission was significantly independent of 5577 on all these dates. September 21/22 and 22/23 were also active nights (6300) at Haute Provence.

At Rapid City, we have also noted cases of a 6300 arc apparently detached from and to the south of the aurora proper. An outstanding example of this was October 22/23, 1958, a night which is currently under study by us.

It is not known whether 6300 arcs occur as a general or as a sporadic fringe on the equatorial side of prominent visual auroras. Barbier [1] has made a statistical study of 6300 auroras at Haute Provence and finds that they occur during 9.8 percent of the observing time. Of 31 nights included in his list, four are listed as showing well-defined "arcs." Thus, distinct arcs occurred during approximately $4/31 \times 9.8 = 1.3$ percent of the observing time. According to Vestine [5], auroras at a latitude (geomagnetic) of 46° are visible 1.5 percent of the time. Thus, it is not unreasonable to hypothesize that isolated 6300 arcs are a *general* fringe feature of auroras. If indeed this is the case, a reexamination is in order of all auroral theories based on the assumption of a narrow, well-defined auroral zone.

9. References

[1] D. Barbier, L'activite aurorale aux basses latitudes, Ann. Geophys. **14,** 334 (1958).

[2] S. Chapman, Polar and tropical aurorae and the isoauroral diagram, Proc. Indian Acad. Sci. **37**, 175 (1953); The aurora in middle and low latitudes, Nature **179**, 7 (1957); Auroral observations in India and Pakistan, Bull. Nat. Inst. Sci. (India) No. 9, 180 (1957).

[3] J. Van Allen, C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958 ϵ, J. Geoph. Research **64,** 271 (1959) and Scientific American (March 1959).

[4] L. R. Megill, Ph. D. dissertation, A study of the atmospheric twilight emissions [OI] 5577 and [OI] 6300, University of Colorado (1959).

[5] E. H. Vestine, The geographic incidence of aurora and magnetic disturbance, northern hemisphere, Terrestrial Magnetism Atmospheric Elec. **49,** 77 (1944).

BOULDER, COLO.

(Paper 63D3–26)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Pattern Synthesis for Slotted-Cylinder Antennas*

James R. Wait and James Householder

(June 11, 1959)

The radiation from a cylinder excited by an array of axial slots is discussed. A procedure for synethesizing a given radiation pattern is developed with particular attention being paid to a Tchebyscheff-type pattern. Specifying the side lobe level and the width of the main beam, the required source distributions are computed for a number of cases. The effect of using a finite number of slot elements to approximate the continuous source distribution is also considered.

1. Introduction

It is the purpose of this communication to discuss the synthesis of radiation patterns for antennas of slotted-cylinder type. The particular model employed is a metallic cylinder of infinite length and perfect conductivity. The cylinder is to be energized by an array of slots of rectangular shape which are oriented in the axial direction and disposed around the circumference. Each slot is to be energized in such a way that the tangential field in the slot has only a transverse component. The radiation field of such a slot on an infinite cylinder is plane polarized. Furthermore, the idealization of an infinitely long cylinder is not overly restrictive since the surface currents excited by the axial slot are very similar to what they would be on a finite cylinder. This conjecture is substantiated by experimental data on the conductance of axial slots on finite cylinders [1].[1]

2. Basic Pattern Function

Choosing a cylindrical coordinate system (ρ, ϕ, z), the cylinder is defined by $\rho=a$. Since attention is only confined to the azimuthal or ϕ behavior, it is only necessary to specify the ϕ variation of the field in the axial slot. For example, if the slot width extends from ϕ_1 to ϕ_2, the azimuthal radiation pattern $M^*(x, \phi)$ was shown to be given by [1]

$$M^*(x,\phi)=\frac{1}{i\pi x}\sum_{m=0}^{\infty}\frac{\epsilon_m e^{im\pi/2}\cos m\phi}{H_m^{(2)\prime}(x)}J_0\left(\frac{\phi_2-\phi_1}{2}m\right), \quad (1)$$

where $H_m^{(2)\prime}(x)$ is the derivative with respect to x of the Hankel function of order m, J_o is the Bessel function of order zero, $k=2\pi/\text{wavelength}$, $x=ka\sin\theta$, $\epsilon_0=1$, $\epsilon_m=2$ $(m\neq 0)$, and θ is the angle subtended by the cylinder or z axis and the direction to the observer (i.e., the usual polar angle in spherical coordinates).

The above form for the pattern $M^*(x, \phi)$ is strictly valid only when the field E_ϕ in the slot is

$$E_\phi(\phi_0)=\frac{V(z_1)}{a\pi\sqrt{(\Delta/2)^2-\phi_0^2}}, \quad (\Delta=\phi_2-\phi_1), \quad (2)$$

where $V(z_1)$ is the voltage across the slot at $z=z_1$. The field approaches infinity as the inverse square root of the distance to the edge of the slot. Such a behavior is characteristic of the field in the vicinity of a perfectly conducting knife edge.

The amplitude and normalized phase Φ [=phase $M^*(x, \phi)-x\cos\phi$] of the quantity $M^*(x, \phi)$ is calculated for a series of values of x and $\phi_2-\phi_1$. These are listed in tables 1 and 2 for x equal to 3 and 5, respectively, and $\phi_2-\phi_1$ varying from 0° to 30°. For the smaller cylinder, it is seen that the effect of widening the slot (for a given voltage) is small. For the larger cylinders, the effect is becoming more significant although if the angular slot width is also less than 10°, the pattern is indistinguishable from that of zero width.

It should be emphasized that these E-plane patterns characterize only the azimuthal dependence of the radiation field for a given value of $x(=ka\sin\theta)$. The elevation patterns or H plane pattern of such an axial slot requires that the longitudinal distribution of voltage along the slot be specified [1].

3. General Array Synthesis

By superimposing the patterns of individual axial slots arranged circumferentially around the cylinder, various forms of patterns can be obtained. The problem sometimes arises that a pattern is specified and the manner of excitation must be determined. This is described as synthesis. Some elegant procedures have been developed for linear arrays and these have been described extensively in the literature [2 through 5]. When the antennas are disposed around circular arcs, the techniques [6] are not so straightforward, particularly if diffraction is involved. An example falling in this category will be considered in what follows.

*The research reported in this paper was sponsored, in part, by the U.S. Air Force Cambridge Research Center under Contract CSO and A 58–40.
[1] Figures in brackets indicate the literature references at the end of this paper.

TABLE 1 [a]

(From reference [7])

$x=3$

$\phi_2-\phi_1$	$\lvert M^*(x,\phi)\rvert$			Φ(deg)		
	0	20	30	0	20	30
$\phi=$ 0	0.959	0.957	0.955	6.9	5.7	4.2
10	.956	.953	.948	7.1	5.9	4.5
20	.946	.937	.926	7.7	6.6	5.2
30	.927	.909	.888	8.4	7.4	6.2
40	.895	.869	.838	8.9	8.1	7.2
50	.859	.824	.783	8.8	8.2	7.7
60	.825	.782	.731	8.0	7.9	8.0
70	.794	.743	.684	7.4	8.0	8.7
80	.747	.691	.626	7.8	8.8	10.2
90	.664	.609	.545	7.6	9.0	11.0
100	.561	.512	.455	3.2	5.0	7.5
110	.482	.445	.395	−7.0	−5.1	−2.0
120	.470	.428	.379	−18.1	−15.4	−11.7
130	.444	.404	.358	−23.5	−20.4	−16.1
140	.349	.320	.283	−28.7	−23.9	−19.2
150	.206	.187	.166	−43.5	−40.0	−35.2
160	.155	.142	.126	−105.0	−101.7	−97.0
170	.259	.237	.211	−137.0	−133.5	−128.6
180	.312	.285	.254	−142.1	−138.5	−133.6

[a] All angles are shown in degrees.

TABLE 2 [a]

(From reference [7])

$x=5$

$\phi_2-\phi_1$	$\lvert M^*(x,\phi)\rvert$				Φ(deg)			
	0	10	20	30	0	10	20	30
$\phi=$ 0	0.981	0.980	0.979	0.979	4.9	4.5	2.8	0.2
10	.977	.976	.970	.960	5.1	4.7	3.0	.5
20	.966	.963	.945	.918	5.1	5.1	3.5	1.2
30	.951	.944	.908	.855	5.8	5.5	4.1	2.3
40	.936	.924	.865	.782	6.1	5.9	5.0	3.8
50	.914	.899	.815	.701	7.0	6.9	6.4	6.2
60	.865	.844	.745	.610	7.9	7.9	8.0	8.7
70	.804	.781	.672	.527	7.0	7.1	7.9	10.1
80	.757	.732	.615	.462	5.5	5.8	7.6	11.6
90	.681	.656	.540	.392	5.0	5.4	8.0	13.7
100	.570	.548	.447	.318	−0.6	0	3.4	10.5
110	.513	.493	.400	.282	−11.3	−10.5	−6.2	2.7
120	.452	.434	.350	.245	−17.5	−16.6	−11.6	−1.3
130	.318	.306	.246	.172	−32.7	−31.8	−26.4	−15.3
140	.288	.277	.224	.158	−68.0	−67.0	−61.4	−49.8
150	.293	.281	.228	.162	−82.8	−81.7	−75.7	−63.5
160	.147	.141	.114	.082	−99.7	−98.8	−92.6	−77.0
170	.142	.137	.112	.080	−205.2	−204.0	−198.0	−185.7
180	.241	.232	.189	.135	−219.3	−218.2	−212.1	−199.6

[a] All angles are shown in degrees.

The cylinder is now to have P axial slots equispaced around a circumference of the cylinder as illustrated in figures 1a and 1b. The individual slot elements are centered at

$$\phi=\phi_p=\frac{2\pi p}{P} \text{ where } p=0, 1, 2 \ldots P-1.$$

The quantity $L(\phi_p)$ is used to describe the relative excitation of each element. Therefore, the resultant azimuthal pattern of the array is

$$\overline{M}(\phi)=\frac{1}{2\pi}\sum_{p=0,1,2\ldots}^{P-1} L(\phi_p)M^*(\phi-\phi_p) \qquad (3a)$$

FIGURE 1a. *The axial slotted-cylinder array.*

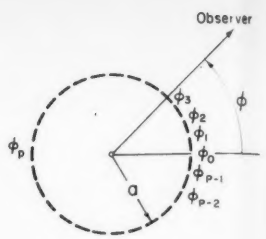


FIGURE 1b. *Cross section of the cylinder.*

where $M^*(\phi-\phi_p)$ is the pattern of the individual elements of the array. (The factor $1/2\pi$ is included for convenience in what follows.) Now when P is sufficiently large, $\overline{M}(\phi)$ can be approximated by an integral so that p is then regarded as a continuous variable. This leads to

$$\overline{M}(\phi)\cong\frac{1}{2\pi}\int_0^{2\pi} L(\phi_p)M^*(\phi-\phi_p)d\phi_p. \qquad (3b)$$

It is now assumed that the excitation function $L(\phi_p)$ is expressible as a Fourier series in the manner

$$L(\phi_p)=\sum_{n=-\infty}^{+\infty} L_n e^{in\phi_p} \qquad (4)$$

where the sum is over all integral values of n (including positive and negative integers). It is also assumed that the width of the individual slots is very small and therefore

$$M^*(\phi-\phi_p)\cong\frac{1}{i\pi x}\sum_{m=-\infty}^{+\infty}\frac{e^{im\pi/2}e^{im(\phi-\phi_p)}}{H_m^{(2)\prime}(x)} \qquad (5)$$

which follows directly from eq(1). Inserting these expressions for $L(\phi_p)$ and $M^*(\phi-\phi_p)$ into eq(3b) leads to

$$\overline{M}(\phi)=\frac{1}{i\pi x}\sum_{m=-\infty}^{+\infty}\frac{L_m e^{im\pi/2}e^{im\phi}}{H_m^{(2)\prime}(x)}. \qquad (6)$$

Having the pattern expressed in this form enables a synthesis procedure to be directly applied. Because of orthogonality

$$L_m=\frac{ix}{2}e^{-im\pi/2}H_m^{(2)\prime}(x)\int_0^{2\pi}\overline{M}(\phi)e^{-im\phi}d\phi \qquad (7)$$

which, when inserted into eq(4), enables $L(\phi_p)$ to be determined and is the excitation required to produce the pattern $\overline{M}(\phi)$. In a formal sense, any pattern $\overline{M}(\phi)$ could be specified, however directive, and a corresponding function $L(\phi_p)$ could be determined. From a practical standpoint, there is a limitation

when $\overline{M}(\phi)$ becomes very directive since then $L(\phi_P)$ varies very rapidly in both amplitude and phase and leads to a complicated procedure for feeding the array. Furthermore, small errors in the excitation of the elements lead to a large degradation of the pattern.

A different approach to the synthesis problem is to utilize the remarkable properties of the Tchebyscheff polynomials. These polynomials are particularly appropriate when the pattern is to have side lobes of specified and equal amplitude. As shown by Dolph [2] and others [3, 4, 5, 6], the resultant design is an optimum one for a linear array of discrete elements. In the case of a circular array, a similar procedure can be adopted as follows:

The Tchebyscheff polynomial of order N is defined by

$$\begin{aligned} T_N(z) &= \cos\ (N \text{ arc cos } z),\ z \leqq 1, \\ &= \cosh\ (N \text{ arc cosh } z),\ z > 1. \end{aligned} \tag{8}$$

In order to utilize the optimum properties of the polynomial, the following transformation is introduced

$$z = a \cos \phi + b,$$

where a and b are constants. Now $\phi = 0$ and $z = z_0$ are to correspond to the direction of the main beam. In fact, as ϕ varies from 0 to 2π, z is to vary from z_0 to -1 back to z_0. Therefore

$$a = \frac{z_0 + 1}{2} \quad \text{and} \quad b = \frac{z_0 - 1}{2}.$$

As shown by Duhamel [5],

$$T_N(a \cos \phi + b) = \sum_{n=0}^{N} C_n^N \cos n\phi, \tag{9}$$

where C_n^N are functions of a and b. For example, when $N = 4$, they are given by

$$\begin{aligned} C_0^4 &= 1 - 8b^2 + 8b^4 + 24a^2b^2 + 3a^4 - 4a^2 \\ C_1^4 &= -16ab + 32ab^3 + 24a^3b \\ C_2^4 &= -4a^2 + 24a^2b^2 + 4a^4 \\ C_3^4 &= 8a^3b \\ C_4^4 &= a^4 \end{aligned} \tag{10}$$

and for $N = 6$, they are given by

$$\begin{aligned} C_0^6 &= -1 + 18b^2 - 48b^4 + 32b^6 + 9a^2 - 144b^2a^2 \\ &\quad + 240b^4a^2 - 18a^4 + 180b^2a^4 + 10a^6 \\ C_1^6 &= 36ba - 192b^3a + 192b^5a - 144ba^3 + 480b^3a^3 + 120ba^5 \\ C_2^6 &= 9a^2 - 144b^2a^2 + 240b^4a^2 - 24a^4 + 240b^2a^4 + 15a^6 \\ C_3^6 &= -48ba^3 + 160b^3a^3 + 60ba^5 \\ C_4^6 &= -6a^4 + 60b^2a^4 + 6a^6 \\ C_5^6 &= 12ba^5 \\ C_6^6 &= a^6. \end{aligned}$$

The next step is to equate the Tchebyscheff pattern to the general form of the slotted cylinder pattern given by eq (6). Thus

$$T_N(a \cos \phi + b) = \frac{1}{i\pi x} \sum_{m=0}^{\infty} \frac{\epsilon_m L_m e^{im\pi/2} \cos m\phi}{H_m^{(2)\prime}(x)}. \tag{11}$$

This enables the coefficients L_m to be specified in terms of the coefficients C_m^N as follows

$$\begin{aligned} \epsilon_m L_m &= i\pi x e^{-im\pi/2} C_m^N H_m^{(2)\prime}(x) \text{ for } m \leqq N, \\ &= 0 \qquad\qquad\qquad\qquad\quad \text{for } m > N. \end{aligned} \tag{12}$$

The corresponding excitation is then given by

$$L(\phi_p) = \sum_{m=0}^{\infty} \epsilon_m L_m \cos m\phi_p. \tag{13}$$

It is thus possible to produce a Tchebyscheff type of pattern from a distribution of axial slots around a circular cylinder. As shown below, either the side-lobe level or the first null can be specified by a proper choice of z_0. For a constant side-lobe level, the beam width is decreased by increasing the order of the polynomial.

The nulls of $T_N(z)$ occur when

$$\text{arc cos } z_k = \left(\frac{2k-1}{2N}\right)\pi, k = 1, 2, 3 \ldots N$$

and the nulls in the ϕ domain are thus

$$\phi_k = \pm \text{arc cos} \left[\frac{\cos\left[(2k-1)\frac{\pi}{2N}\right] - b}{a}\right], k = 1, 2 \ldots N. \tag{14}$$

On the other hand, the center of the lobes occur at $\phi = \phi_k'$ which is a solution of

$$dT_N(a \cos \phi + b)/d\phi = 0.$$

These are given by

$$\phi_k' = \pm \text{arc cos} \left[\frac{\cos \frac{k\pi}{N} - b}{a}\right]. \tag{15}$$

When the ratio of the side lobe to the main beam is B then

$$T_N(z_0) = B \tag{16}$$

since

$$T_N(\pm 1) = 1.$$

The quantity B thus specifies z_o which in turn determines a, b, ϕ_k and ϕ_k' for a given value of N. The required excitation $L(\phi_p)$ actually is a continuous function. From a practical standpoint, the number of elements must be finite. It would be expected, however, if the separation between the elements is small compared to the wavelength, the actual pat-

tern would not differ appreciably from the Tchebyscheff pattern. This conjecture is verified by actually computing the pattern for a finite number of equispaced slots disposed around the cylinder. For example, the polynomial $T_N(a\cos\phi+b)$ is to be compared with the function

$$\overline{M}(\phi)=\frac{1}{P}\sum_{p=0,1,2\ldots}^{P-1} L(\phi_p)M^*(\phi-\phi_p) \qquad (17)$$

with P being a finite integer. In this case, the spacing between the axial slots is $2\pi a/P$. In spite of the fact that the spacing is finite, the $\overline{M}(\phi)$ pattern is quite close to the ideal T_N pattern as indicated in the following results.

4. Presentation of Numerical Results

The theoretical basis of the synthesis procedure has been outlined in the previous section. Here numerical results are presented for certain specific cases which illustrate the interrelation between the parameters of the problem.

The basic patterns used are illustrated in figure 2a and 2b. They are derived from the T_4 and T_6 polynomials, respectively. The quantity B, which is the (voltage) ratio of the major lobe to the minor lobe, is taken as 5, 10, 15, and 20 for the two cases. It is seen that for a given polynomial the beam is broadened as the side-lobe level is reduced. Consequently, if for a given side-lobe level, one wishes to

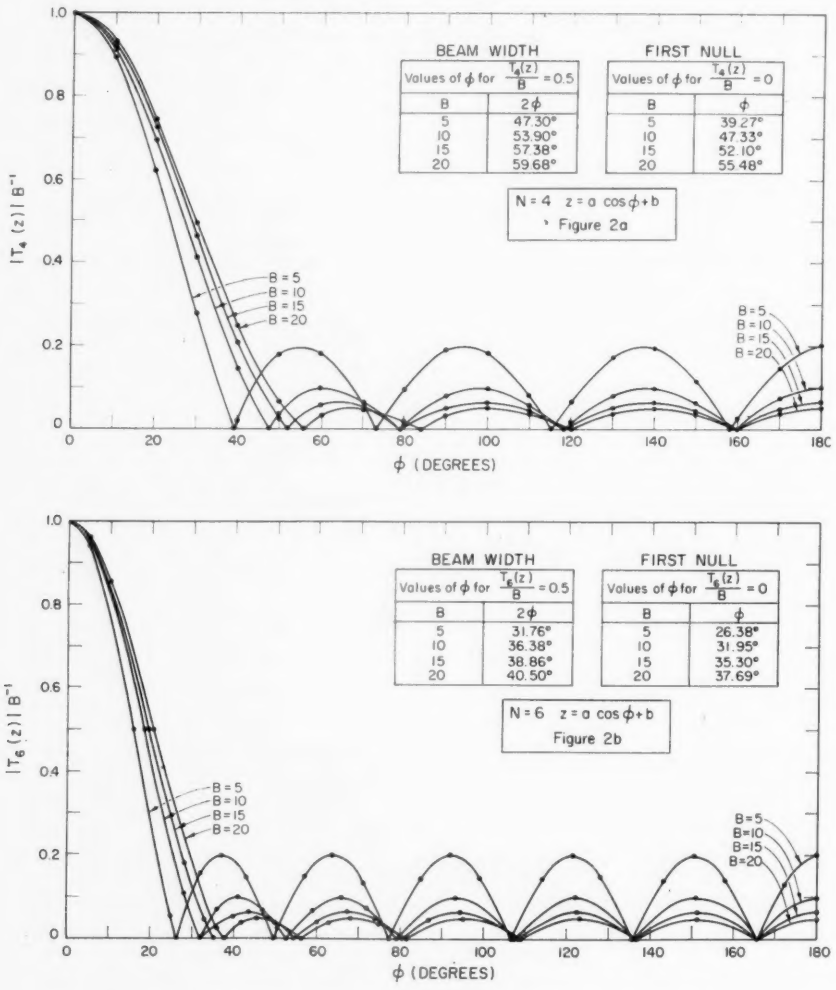
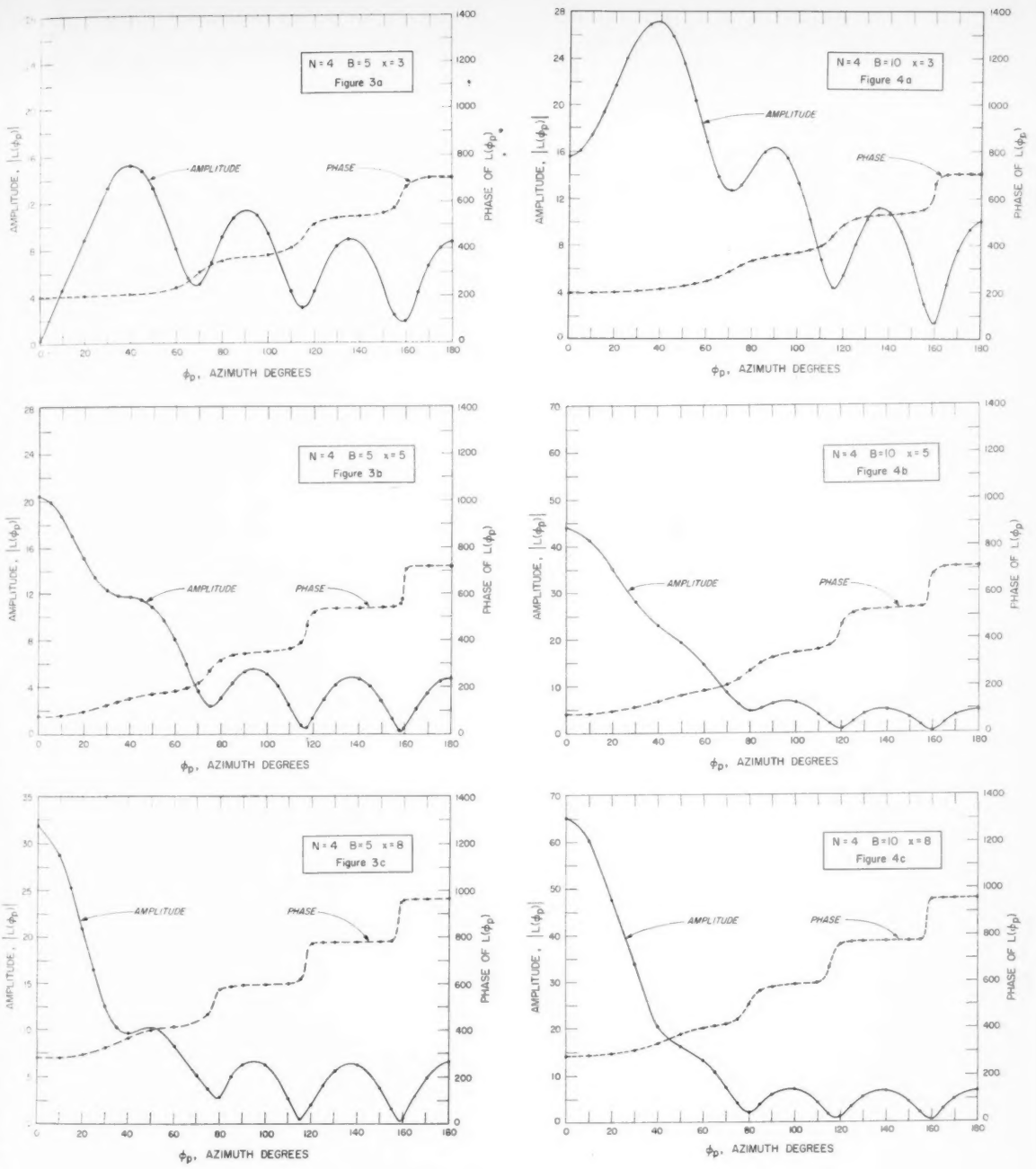
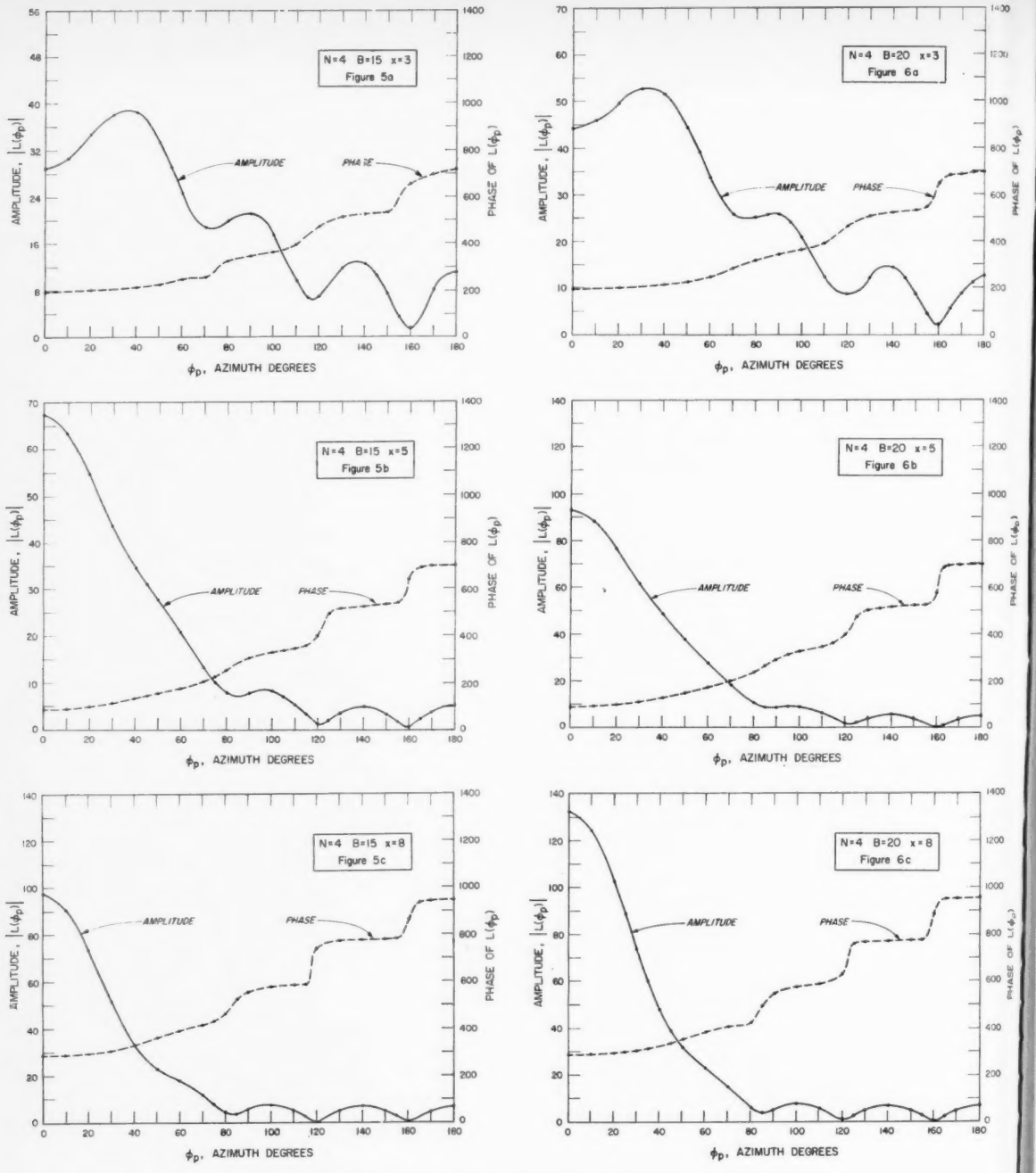


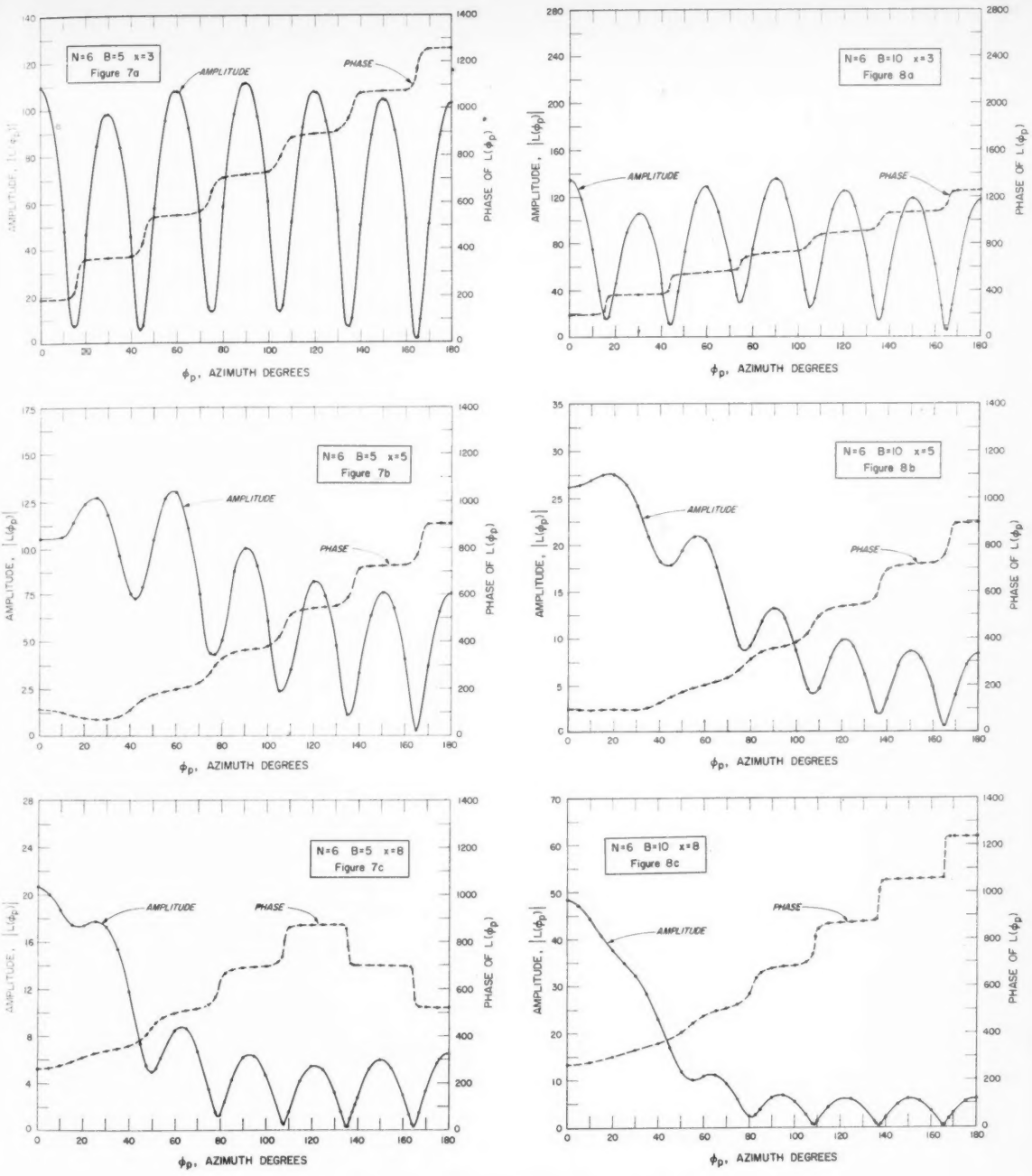
FIGURE 2. *Optimum patterns based on Tchebyscheff polynomials* $T_N(z)$.


N = 4 B = 5 x = 3
Figure 3a
N = 4 B = 10 x = 3
Figure 4a
N = 4 B = 5 x = 5
Figure 3b
N = 4 B = 10 x = 5
Figure 4b
N = 4 B = 5 x = 8
Figure 3c
N = 4 B = 10 x = 8
Figure 4c
AMPLITUDE
PHASE
AMPLITUDE, |L(φp)|
PHASE OF L(φp)
φp, AZIMUTH DEGREES

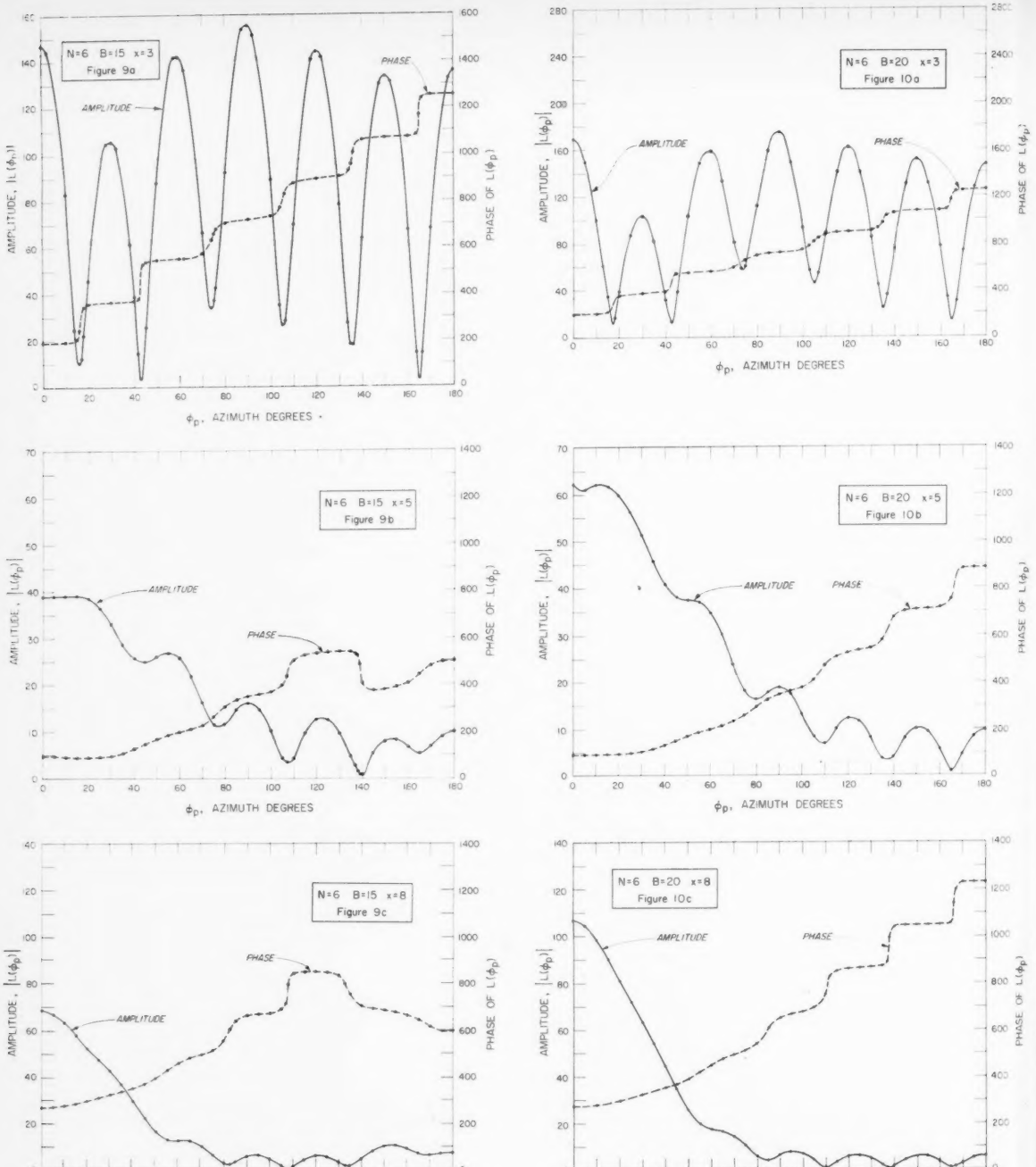
FIGURES 3 and 4. *Excitation required for a Tchebyscheff pattern.*



FIGURES 5 and 6. *Excitation required for a Tchebyscheff pattern.*



FIGURES 7 and 8. *Excitation required for a Tchebyscheff pattern.*



FIGURES 9 and 10. *Excitation required for a Tchebyscheff pattern*

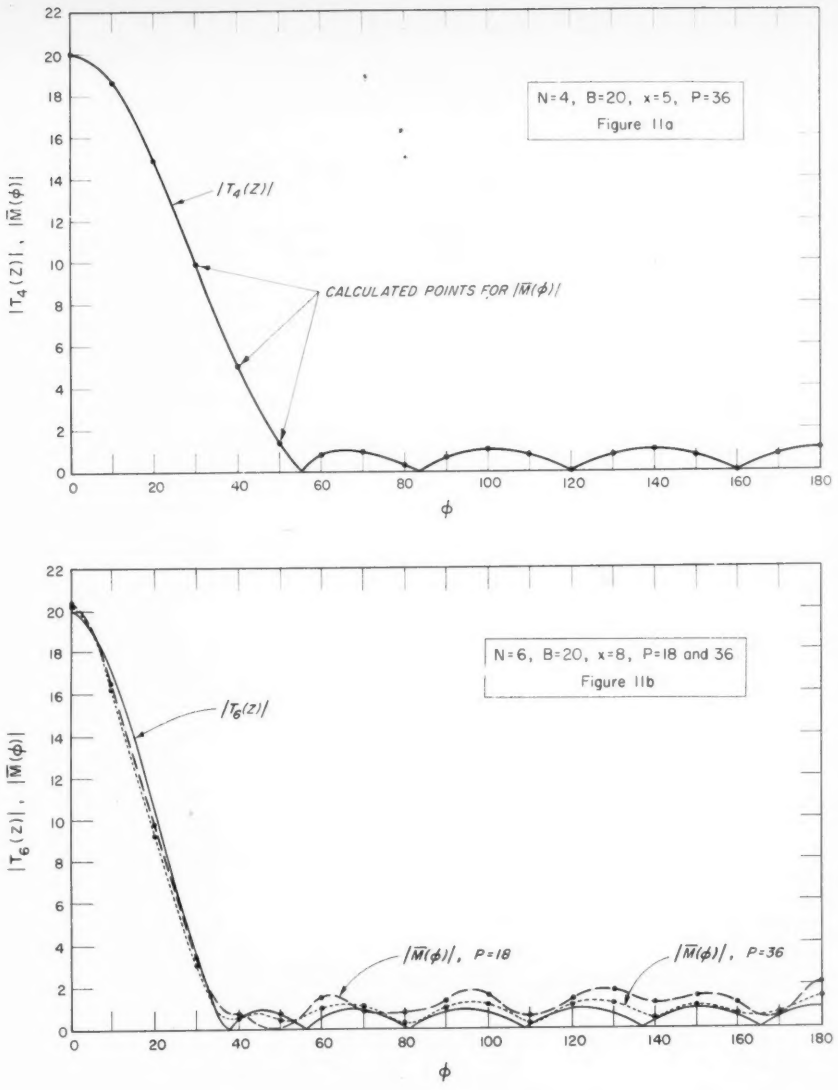


FIGURE 11. $|\overline{M}(\phi)|$ *compared with* $|T_N(z)|$.

reduce the beam width, a higher order polynomial must be employed. This fact is illustrated by noting that the beam width is narrower for the T_6 polynomial than for the T_4 polynomial for a given value of B.

The curves shown in figure 3a, 3b, 3c to figure 10a, 10b, 10c, inclusive, are plots of the amplitude and phase of the excitation function $L(\phi_p)$. The abscissas are the azimuthal coordinates ϕ_p expressed in degrees. On each figure is listed the revelant value of the order N of the Tchebyscheff polynomial, the beam ratio B and cylinder parameter x.

It should be noted that the quantity x, which is $ka \sin \theta$, is fixed for each set of curves. In the equatorial plane (i.e., $\theta=90°$) x becomes equal to ka, which is the circumference of the cylinder in wavelengths.

In figure 7a (for $x=3$) it is noted that the excitation amplitude is varying in a very pronounced manner. Here, although the diameter of the cylinder is only about one-half wavelength, the main beam width is less than 32°. In figure 7b (for $x=5$) and figure 7c (for $x=8$), the oscillatory nature of the

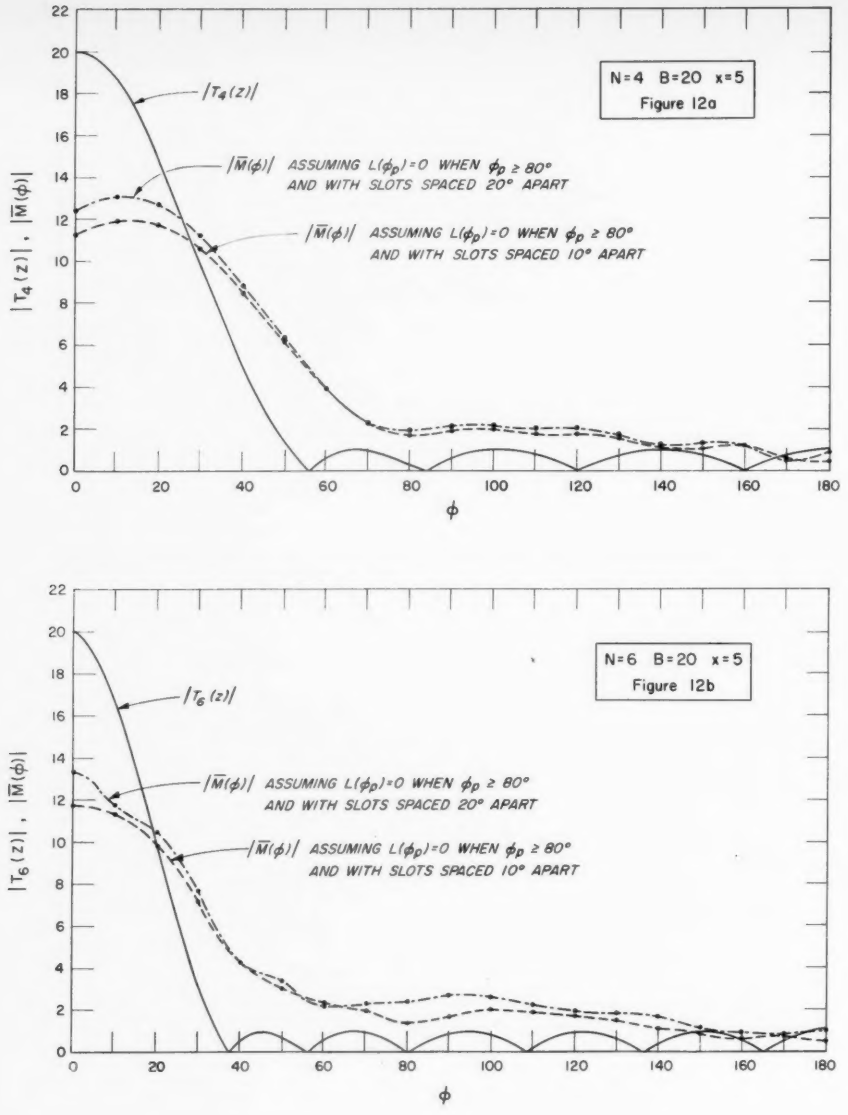

N=4 B=20 x=5
Figure 12a
$|T_4(z)|$
$|\overline{M}(\phi)|$ ASSUMING $L(\phi_p)=0$ WHEN $\phi_p \geq 80°$ AND WITH SLOTS SPACED 20° APART
$|\overline{M}(\phi)|$ ASSUMING $L(\phi_p)=0$ WHEN $\phi_p \geq 80°$ AND WITH SLOTS SPACED 10° APART
$|T_4(z)|$, $|\overline{M}(\phi)|$
ϕ
N=6 B=20 x=5
Figure 12b
$|T_6(z)|$
$|\overline{M}(\phi)|$ ASSUMING $L(\phi_p)=0$ WHEN $\phi_p \geq 80°$ AND WITH SLOTS SPACED 20° APART
$|\overline{M}(\phi)|$ ASSUMING $L(\phi_p)=0$ WHEN $\phi_p \geq 80°$ AND WITH SLOTS SPACED 10° APART
$|T_6(z)|$, $|\overline{M}(\phi)|$
ϕ

FIGURE 12. *Pattern comparison for various assumptions.*

excitation is less evident. The following sequence of figures for the larger B values show a similar trend.

In nearly every case shown in figure 3a to figure 10c, the phase of the excitation increases as ϕ_p ranges from 0° to 180°. When the amplitude is rapidly varying such as in figure 7a the phase increases almost abruptly by π radians (or 180°) near each minimum in the amplitude. There are several seemingly anomalous cases (fig. 7c, for example) where the π radian phase jump is down rather than up.

5. Effect of Using Discrete Sources

The excitation functions presented above are really descriptions of continuous source distributions. In a practical scheme it is usually necessary to approximate these by an array of discrete sources. As mentioned above, this is to be accomplished by an array of thin axial slots disposed uniformly around the circumference of the cylinder. In figure 11a is shown the pattern $|\overline{M}(\phi)|$ calculated from eq

(17) using 36 slots to approximate the continuous source distribution given by figure 6b. These calculated points lie virtually on top of the ideal Tchebyscheff (T_4) pattern. The apparent coincidence of the calculated and the ideal patterns is not surprising since the separation between the discrete sources is only $\frac{1}{12}$ of a wavelength, and thus the excitation is virtually a continuous one. A similar set of calculations are shown in figure 11b where the source distribution of figure 10c is approximated by both 18 and 36 discrete elements. There now appears to be some notable departures from the ideal Tchebyscheff pattern, particularly in the case for 18 elements where the side lobes are increased. In both cases, however, the main beam is preserved. In these two cases, the separation between the elements is $\frac{2}{9}$ and $\frac{1}{9}$ of a wavelength, so the degradation of pattern is not unexpected.

On examining some of the excitation functions (for example, fig. 6b), it is seen that on the rear half of the cylinder, the amplitude $|L(\phi_p)|$ is less than 10 percent of the maximum at $\phi_p=0$. As a matter of interest, the function $\overline{M}(\phi)$ was computed using eq (17) again, but now approximating $L(\phi_p)$ by discrete sources located only on a front portion of the cylinder (i. e., $|\phi_p| < 80°$). Such results are shown in figure 12a for angular intervals between the slots of 10° and 20°. The computed pattern $\overline{M}(\phi)$ bears only a slight resemblance to the ideal T_4 pattern which would be closely simulated if the slots extended right around to the back of the cylinder. Similar calculations are shown in figure 12b where the excitation function shown in figure 10c is approximated by discrete sources on the front portion of the cylinder. Here again the pattern is considerably degraded from the ideal T_6 pattern although the increase of the side-lobe level may be tolerated for certain applications.

6. Concluding Remarks

It can be seen from the above curves that, in principle, any pattern may be obtained if the excitation function or source distribution is specified in a certain way. Unfortunately, if narrow main beams are desired, the cylinder must be sufficiently large, otherwise very complicated excitation functions are required.

It appears that usually the continuous excitation or source function $L(\phi_p)$ can be approximated by discrete sources (i.e., narrow axial slots) if the separation is less than about $\lambda/8$. However, the rapidly varying excitation functions associated with small cylinders and narrow beams would require smaller spacing between the source elements.

7. References

[1] J. R. Wait, Electromagnetic radiation from cylindrical structures (Pergamon Press, 1959); and J. R. Wait, Radiation characteristics of axial slots on a conducting cylinder, Wireless Eng. **32** 316 (1955).

[2] C. L. Dolph, A current distribution for broadside arrays which optimizes the relationship between beam width and side lobe level, Proc. IRE **34** 335 (1946).

[3] H. J. Riblet, Discussion of Dolph's Paper, Proc. IRE **35** 489 (1947).

[4] G. Sinclair and F. V. Cairns, Optimum patterns for arrays of non-isotropic sources, IRE Trans. PGAP–1, 50 (1952).

[5] R. H. Duhamel, Optimum patterns for end-fire arrays, Proc. IRE **41** 652 (1953).

[6] R. H. Duhamel, Pattern synthesis for antenna arrays, Tech. Rept. No. 16, Elec. Eng. Research Lab. Univ. of Ill. (May 1952).

[7] J. R. Wait and P. Wheelon, Further numerical results for the patterns of slotted-cylinder antennas, Rept. on Contract AFO4 (645)–56 prepared by P.E.C. Corp. under the direction of D. G. Burkhard for the Martin Co. Denver Div. (June 1958).

Note added in proof:

A closely related problem has been studied by Adolf Giger in *On the Construction of Antenna Arrays with Prescribed Radiation Patterns*, doctoral thesis, Technischen Hochschule, Zurich, 1956. This was brought to our attention recently by Prof. Franz Tank.

Boulder, Colo.

(Paper 63D3–27)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Central Radio Propagation Laboratory Exponential Reference Atmosphere[1]

B. R. Bean and G. D. Thayer

(June 11, 1959)

The background and development of an exponential model of atmospheric radio refractivity, the "C.R.P.L. Exponential Reference Atmosphere" is outlined. A set of ray tracings for the model is presented in the form of tables of refraction variables for the complete range of observed values of surface refractive index. A detailed analysis of the accuracy of the ray tracing and tabulation methods is made for these tables. The variables are presented as numbers between one and ten multiplied by the appropriate power of ten, thus maintaining a maximum number of significant figures. The tables may be used for the solution of practical refraction problems involving elevation angle errors, range errors, and similar quantities.

1. Introduction

Some exponential models of atmospheric radio refractivity have been introduced recently to replace the standard 4/3 earth treatment for applications involving long-range transmission paths, forward scatter predictions, or radar tracking at high altitudes [1, 2, 3, 4, 5, 6,].[2] One of these models, developed at the Central Radio Propagation Laboratory and called the "C.R.P.L. Exponential Reference Atmosphere" has been adopted for use by the National Bureau of Standards in predictions of refraction phenomena.

In the Monograph the background of the exponential reference atmosphere is outlined and a set of radio ray tracings presented. The ray tracings may be used for the solution of many problems involving refraction.

2. Background

In dealing with refraction problems it is convenient to use the radio refractivity defined by

$$N \equiv (n-1)\times 10^6$$

rather than the radio refractive index, n. The basic predictor of refraction effects used in the exponential reference atmosphere is the surface value of refractivity at the transmitting point, N_s. The reasons for this become clear upon examination of the normal, or average, vertical N-structure of the atmosphere.

There is an observed high correlation between the mean value of ΔN, the drop in N from the surface to a height of 1 km, and mean value of N_s. This is expressed by

$$-\Delta N = 7.32 \exp\{0.005577\, N_s\} \qquad (1)$$

which has been derived from 6 yr means for 45 U.S weather stations representing all types of climatic conditions encountered in the United States. Thus it is evident that on the average the decay of N with height may be expressed as a function of the surface value of N, at least over the lower portion of the earth's atmosphere which is the critical portion when considering refraction effects.

It is also found that the average refractivity structure of the atmosphere follows a nearly exponential decay with height, the value of N becoming insignificantly small at altitudes over 50 km.

For these reasons the exponential reference atmosphere was defined as that family of exponential refractivity profiles where the rate of N decay with height is uniquely specified by the value of N_s in accordance with eq (1). The exponential reference atmosphere is thus represented by the following set of equations:

$$N = N_s \exp\{-c_e h\}$$

$$c_e = \ln\{N_s/(N_s+\Delta N)\}$$

with ΔN determined from (1) above.

It is found that several important advantages accrue from this particular choice of model, among which are:

1. The N-profiles involved prove to be a fairly good representation of average N-structure over the first 3 km above the surface, a critical range for refraction effects at low angles to the horizon.

2. The resulting relationship between the N-gradient near the surface and N_s is found to be very close to the observed average relationship being within ± 3 percent over the normally encountered range of N_s (from 300 to 365 for sea level, temperate zone stations).

3. The resulting profiles and their associated gradients are continuous functions of height and are in reasonably good agreement with the actual

[1] Digest of NBS Monograph (in press).
[2] Figures in brackets indicate the literature references at the end of this paper.

observed refractivity at all heights below the ionosphere.

An illustration of the effectiveness of the exponential reference atmosphere in predicting average refraction effects is given by figure 1, where values of angular ray bending are plotted against height for rays starting at zero elevation angle. The agreement between the average bending for 22 climatically diverse conditions and the average of the values predicted for each climatic condition by the exponential reference atmosphere is seen to be quite good, particularly as regards the form of the τ versus height curve. Note, however, that the bending versus height predicted by the 4/3 earth model is markedly in disagreement with the observed values for most heights.

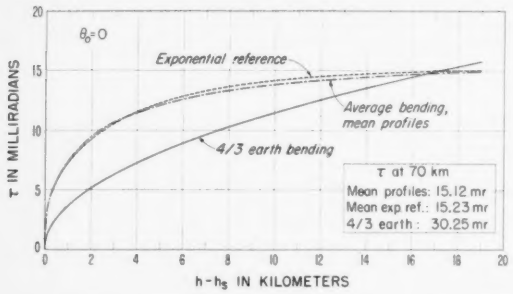


FIGURE 1. *Comparisons of long-term mean bending with (a) the average of the values predicted by the appropriate exponential reference atmospheres and with (b) those predicted for the 4/3 earth's atmosphere.*

The preceding material is intended only as a background for the development of the exponential reference atmosphere; a more detailed treatment may be found in reference [5].

The remainder of the Monograph is devoted to a detailed analysis of the ray-tracing theory used to calculate refraction variables for the exponential reference atmosphere, together with an analysis of the actual method of calculation. The extensive tables of refraction variables are presented in an appendix to the Monograph.

3. Tables of the Exponential Reference Atmosphere

The tables of refraction variables for the "Exponential Reference Atmosphere" are presented in the form of ray tracings for profiles with N_s=200.0, 252.9, 289.0, 313.0, 344.5, 377.2, 404.9, and 450.0. The uneven values of N_s (252.9, 344.5, 377.2, 404.9) were chosen to have even values of ΔN (30.0, 50.0, 60.0, 70.0). The ray tracings are presented for initial elevation angles of 0, 0.5, 1, 2, 4, 8, 15, 30, 65, 100, 200, 400, and 900 milliradians, the refraction variables being tabulated at constant heights (of the ray path above the surface) consisting of 0.01, 0.02, 0.05, 0.1, 0.2, 0.305 (1,000 ft), 0.5, 0.7, 1.0, 1.524 (3,000 ft), 2.0, 3.048 (10,000 ft), 5.0, 7.0, 10.0, 20.0, 30.48 (100,000 ft), 50.0, 70.0, 90.0, 110.0, 225.0, 350.0, and 475.0 km. Thus there are 104 ray tracings, each tabulated at 24 heights.

The following refraction variables are tabulated at each height: the local elevation angle, θ; the distance, d, subtended on the earth's surface by the ray path from a height of zero to the tabulated height; the angular ray bending, τ; the elevation angle error, ϵ; the slant range, R_o, (straight line distance from origin to the point on the ray path being considered); the range error, ΔR, (the difference between the geometric length of the curved ray path, $R=\int dR$, and the slant range); the radar range error, ΔR_e, (the difference between the electrical ray path, $R_e=\int ndR$, and the slant range); Δh, the difference between the tabulated height and the height of a ray calculated with the 4/3 earth assumption at the same distance along the earth's surface; and $\Delta\theta$, the difference between θ for the exponential reference atmosphere and θ for the 4/3 earth ray at the same height. All entries are given in a form of floating point notation where all numbers are between 1 and 10 multiplied by the appropriate power of 10, and the number $\pm N.N\times10^{\pm m}$ is printed as $\pm N.N\pm m$, where the second sign is the sign of the exponent of ten.

Four refraction quantities which are not listed can be easily obtained from the tables: the angular distance subtended by the ray path at the center of the earth, $\phi=(d/r_o)$ where r_o is the assumed radius of the earth at the ray path origin; the difference between the elevation angle of the ray path and the elevation angle of the slant range "path", $\delta=\tau-\epsilon$; the geometric ray path length, $R=R_o+\Delta R$; the electrical ray path length, $R_e=R_o+\Delta R_e$.

The elevation angles given in the tables are subject to a maximum error of ±0.00005 mr; the values of τ and ϵ are subject to a maximum error of ±0.001 mr; the values of the different ranges and range errors are subject to a maximum error of ±5 m, in particular the range errors are assumed to have a maximum error of ±0.5 percent or ±0.1 m, whichever is larger.

An example of the refraction tables contained in the appendix to the Monograph is given in the following table for the N_s=313 profile, θ_o=0 (313 is the average value of N_s determined from nearly two million weather observations taken at 45 different weather stations throughout the United States).

As an example of the calculation of the four unlisted quantities previously mentioned, one obtains for the height of 1 km:

$$\phi = \frac{d}{r_0} = \frac{133.073}{6373.2284} = 0.0208800 \text{ radian}$$

$$\delta = \tau - \epsilon = 5.71671 - 2.9263 = 2.7904 \text{ mr}$$

$$R = R_o + \Delta R = 133.08495 + 0.00018 = 133.08513 \text{ km}$$

$$R_e = R_o + \Delta R_e = 133.08495 + 0.03993 = 133.12488 \text{ km}.$$

TABLE 1. *Example of tabulated refraction variables for $\theta_0=0$ and $N_s=313$*

Surface N	Init theta	Height	Delta H	Theta	Delta theta	Distance	Tau	Error angle	Slant range	Delta R	Delta $R-E$
313.0	0.0	0.010	−5.1424−4	1.49613	7.6990−2	1.33678 1	6.01363−1	3.0068−1	1.3367819 1	0.000	4.181-3
313.0	.0	.020	−1.0273-3	2.11617	1.0849−1	1.89044 1	8.50046−1	4.2516−1	1.8904449 1	1.000−6	5.912-3
313.0	.0	.050	−2.5545-3	3.34743	1.6960−1	2.98865 1	1.34194	6.7171−1	2.9886667 1	3.000-6	9.334-3
313.0	.0	.100	−5.0557−3	4.73735	2.3521−1	4.22550 1	1.89273	9.4849−1	4.2255465 1	6.000−6	1.317−2
313.0	.0	.200	−9.9458−3	6.70910	3.2035−1	5.97338 1	2.66351	1.3382	5.9734900 1	2.000-5	1.854-2
313.0	.0	.305	−1.4835−2	8.29831	3.7788−1	7.37272 1	3.26995	1.6474	7.3729190 1	3.500−5	2.278−2
313.0	.0	.500	−2.3345−2	1.06519 1	4.4631−1	9.43092 1	4.14574	2.0975	9.4313420 1	6.600−5	2.890−2
313.0	.0	.760	−3.1325−2	1.26373 1	4.8205−1	1.11483 2	4.85517	2.4678	1.1149053 2	1.000−4	3.387−2
313.0	.0	1.000	−4.2035−2	1.51633 1	4.9670−1	1.33073 2	5.71671	2.9263	1.3308495 2	1.800−4	3.993−2
313.0	.0	1.524	−5.6544−2	1.88405 1	4.4550−1	1.63885 2	6.87418	3.5600	1.6390785 2	3.100−4	4.812−2
313.0	.0	2.000	−6.5784−2	2.17029 1	3.4521−1	1.87356 2	7.69459	4.0267	1.8739027 2	4.600−4	5.398−2
313.0	.0	3.048	−7.3319−2	2.70927 1	7.6253−3	2.30289 2	9.04109	4.8367	2.3035214 2	7.800−4	6.369−2
313.0	.0	5.000	−4.6267−2	3.53155 1	−8.6362−1	2.92785 2	1.06243 1	5.9039	2.9291690 2	1.400-3	7.539−2
313.0	.0	7.000	2.4347−2	4.23594 1	−1.8936	3.44203 2	1.16982 1	6.6859	3.4442178 2	2.000-3	8.284−2
313.0	.0	10.000	1.9795−1	5.15471 1	−3.5381	4.07963 2	1.24648 1	7.5264	4.0833598 2	2.790-3	8.960−2
313.0	.0	20.000	1.1676	7.53020 1	−8.7861	5.65227 2	1.33857 1	9.0531	5.6628165 2	4.460−3	9.801−2
313.0	.0	30.480	2.5569	9.43972 1	−1.3439 1	6.87953 2	1.35470 1	9.8441	6.8993595 2	5.470−3	1.003−1
313.0	.0	50.000	5.6578	1.22337 2	−2.0398 1	8.66238 2	1.35807 1	1.0615 1	8.7039820 2	6.520−3	1.017−1
313.0	.0	70.000	9.2212	1.45417 2	−2.6167 1	1.01334 3	1.35824 1	1.1051 1	1.0202252 3	7.100−3	1.024−1
313.0	.0	90.000	1.3025 1	1.65210 2	−3.1115 1	1.13949 3	1.35825 1	1.1335 1	1.1495030 3	7.590−3	1.028−1
313.0	.0	110.000	1.7003 1	1.82787 2	−3.5509 1	1.25151 3	1.35825 1	1.1540 1	1.2650293 3	7.800−3	1.031−1
313.0	.0	250.000	4.7697 1	2.74520 2	−5.8443 1	1.83614 3	1.35825 1	1.2210 1	1.8820256 3	8.600−3	1.038−1
313.0	.0	350.000	7.1613 1	3.23154 2	−7.0601 1	2.14610 3	1.35825 1	1.2419 1	2.2215878 3	8.900−3	1.042−1
313.0	.0	475.000	1.0321 2	3.73844 2	−8.3274 1	2.46915 3	1.35825 1	1.2584 1	2.5875153 3	9.000−3	1.042−1

4. References

[1] M. Schulkin, Average radio-ray refraction in the lower atmosphere, Proc. I.R.E. **40,** 554 (1952).

[2] L. J. Anderson, Tropospheric bending of radio waves, Trans. American Geophys. Union, **39,** 208 (1958).

[3] J. R. Bauer, W. C. Mason, F. A. Wilson, Radio refraction in a cool exponential atmosphere. Lincoln Labs. Tech. Rept. No. 186 (Aug. 27, 1958).

[4] G. Millington, Propagation at great heights in the atmosphere, Marconi Rev. **21,** 143 (Fourth Quart, 1958).

[5] B. R. Bean and G. D. Thayer, On models of the atmospheric radio refractive index, Proc. I.R.E. **47,** 740 (1959).

[6] B. R. Bean, Atmospheric bending of radio waves, paper presented at Intern. Conf. on Electromagnetic Wave Propagation, Liege, Belgium, Oct. 1958; to be published by the Academic Press.

BOULDER, COLO.

(Paper 63D3-28)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

Excitation Mechanisms of the Oxygen 5577 Emission in the Upper Atmosphere[1]

E. Tandberg-Hanssen and F. E. Roach

(June 18, 1959)

Possible excitation mechanisms for the green 5577 emission are considered in the light of recent data on the dynamics of the upper atmosphere. Photochemical reactions as affected by mass motions as well as excitation directly due to the mass motions are analyzed. It is concluded that either or both mechanisms could probably account for the observed emission.

1. Introduction

It is now well established that the bulk of the λ5577 airglow is generated in a fairly thin layer of the atmosphere at a height of about 100 km (Heppner, Stolarik, and Meredith [1];[2] Roach, Megill, Rees, and Marovich [2]; and Tousey [3]). Furthermore, it is known that the airglow[3] is not emitted in a uniform static layer, but that the emission is concentrated in cells that move relative to the earth's surface (Roach, Tandberg-Hanssen, and Megill [4, 5].

A general theory of the 5577 night airglow which includes both the microscopic (excitation) mechanism and the macroscopic (dynamical cells) features is required, and in this paper two different possibilities are examined.

The principal observational facts any theory will have to explain are summarized:

(1) The average night intensity of the 5577 emission corresponds to approximately 2×10^8 transitions/cm² column sec. For a night of roughly 10 hr duration, $2\times10^8\times3.6\times10^4=7\times10^{12}$ transitions/cm² column are consequently needed.

(2) The airglow intensity varies during one night sometimes by as much as a factor 5, and the maximum comes generally *not* in the beginning of the night (Roach, Tandberg-Hanssen, and Megill [4]).

(3) The emission is not uniform over the sky, but appears in "cells" having horizontal motions of the order of 100 m/sec.

The energy necessary to excite the ^{1}S level of the oxygen atom (from which the atom may jump to the ^{1}D state and emit the green line) is 4.2 ev (32,000° K). The temperature in the 100-km region is less than 300° K which corresponds to 0.04 ev, so it is impossible that any significant excitation can be brought about by the thermal motions of the particles. A static theory must therefore draw on exothermic photochemical reactions which are considered in section 2. Similarly, a dynamic theory must invoke environmental effects whereby energy is given perferentially to selected particles, as discussed in section 4.

2. Photochemical Reactions

Above 100 km most atmospheric oxygen is in the atomic form, and according to Seaton [6], the most important recombination mechanism will be of the 3-body type, as proposed by Chapman [7]. This conclusion is adopted here and the Chapman reaction is considered as the most favorable photochemical excitation mechanism for the green oxygen line. The reaction may be written

$$0+0+0\rightarrow0(^1\mathrm{S})+0_2. \tag{1}$$

The reaction coefficient k for reaction (1) may probably be written (Bates [8])

$$k=5\times10^{-34}\,T^{\frac{1}{2}}\ \mathrm{cm^6/sec}. \tag{2}$$

There is, however, no direct measurement of k and the dependence $T^{\frac{1}{2}}$ is purely conventional. The yield of oxygen atoms excited to the ^{1}S state is given by

$$Y(^1\mathrm{S})=pk[n(0)]^3\ \mathrm{cm^{-3}\ sec^{-1}} \tag{3}$$

where p is an "efficiency factor," giving the fraction of 3-body collisions (1) that lead to excitation of the ^{1}S state, and $n(0)$ is the number density of atomic oxygen.

Since the distribution of excited atoms and molecules over the different energy levels is not known, the value of p is not known. In these numerical calculations, the simplifying assumption is made that $p=1/n$ where n is the number of different types of reactions (1). Taking $p=0.2$ may be a large efficiency factor, but it is at present probably the best.

---

[1] High Altitude Observatory, University of Colorado and National Bureau of Standards, Boulder Laboratories.
[2] Figures in brackets indicate the literature references at the end of this paper.
[3] The following is restricted uniquely to the green emission, λ5577.

Because of the 3-body collision, the airglow emission will be critically height dependent if the distribution of atomic oxygen is at all sharply peaked at any height. If τ_c is the lifetime of the ^1S level for collisional deactivation and τ_r is the radiative lifetime, then the number of photons emitted at any level by metastable oxygen atoms by (1) is proportional (Bates and Massey [9]) to

$$\frac{[n(0)]^3\tau_c}{\tau_c+\tau_r}. \tag{4}$$

Detailed calculations show (Nicolet [10]) that $n(0)$ is peaked near 100 km. According to (2) one should hence expect a pronounced maximum in the green emission near 100 km, and this is in agreement with observations. The number of atoms formed per square centimeter column per second in the ^1S state is then given by

$$N(^1S)=\int_0^\infty Y(^1S)\,dz=p\int_0^\infty k[n(0)]^3dz. \tag{5a}$$

To take into account the actual distribution of $n(0)$ with height may be quite complicated. On the other hand, it is known that it is sharply peaked at $z=100$ km, so most of the contribution from the integral in eq (5) comes from that height. The distribution can hence probably be well approximated with a cosine function or an exponential. An exponential model with a peak value of $n(0)$, taken from Nicolet's [10] work, is used here. Putting

$$n(0)=n_{\max}(0)e^{-z/H}$$

where H is the scale height at 100 km, we find

$$N(^1S)=pk[n_{\max}(0)]^3\int_0^\infty e^{-3z/H}dz$$

$$=1/3K[n_{\max}(0)]^3H=1/3\,Y_0[(^1S)H;K=pk. \tag{5b}$$

Now, according to Garstang [11], the radiative lifetime τ_r for the ^1S state is 0.74 sec, and it is conceivable that collisional deactivation could take place before the $N(^1S)$ atoms jumped to a lower state and emitted the green line. Calling the probability of deactivation d_3, the number of transitions $^1S\rightarrow{}^1D$ is given by

$$N(5577)=1/3\,Y_0(^1S)H(1-d_3)$$
$$\text{transitions/cm}^2\text{ column sec.} \tag{6}$$

The different quantities entering eq (6) are discussed in order to see what numerical values for the green intensity are to be expected from the Chapman reaction.

(a) *Reaction coefficient, k.*

There is no evidence of the effectiveness of the Chapman reaction (see Bates [8]) so the best is to suppose that it is of the same order of magnitude as for other reactions of this type.

For $T=200°$ K, eq (2) gives $k\approx 7\times10^{-33}$

(b) *Density of atomic oxygen, n (0).*

The best estimates (see Nicolet [10]) indicate that $n_{\max}(0)=1.5\text{x}10^{12}$ cm^{-3} between 90 and 105 km, and the density drops off sharply below and above this region.

(c) *Scale height, H.*

The hydrostatic scale height $H=kT/mg$ at this height in the atmosphere is $H=6.5$ km (U.S. Extension to the ICAO Standard Atmosphere [12]).

(d) *Probability of deactivation, d_3.*

The problem of deactivation of the ^1S level has mainly been considered in connection with the intensity ratio of the red to the green emission under auroral conditions. In view of the comparatively short radiative lifetime of the ^1S level, it has generally been assumed that collisional deactivation of this level may be neglected, i.e., $d_3=0$ (see Seaton [13]).

Whereas deactivation by electron collisions may be of importance for bright auroral displays (Seaton [13]), the electron concentrations required seem to be too high for the process to be of importance under ordinary airglow conditions. The rate of recombination ($0+0\rightarrow0_2$) is so slow that there is a negligible decrease in the atomic concentration during the night (Bates and Massey [9]).

In view of the uncertainties in the data, a deactivation of the ^1S level cannot be ruled out. However, it is felt here that there is no obvious process by which the ^1S level will be collisionally depopulated. The factor $(1-d_3)$ is therefore probably close to unity, and this value will be adopted here.

Numerical values for the quantities in (6) will now be put in to see what intensities are to be expected for the green airglow according to the 3-body collision theory. Figure 1 shows the intensities Q in ray-

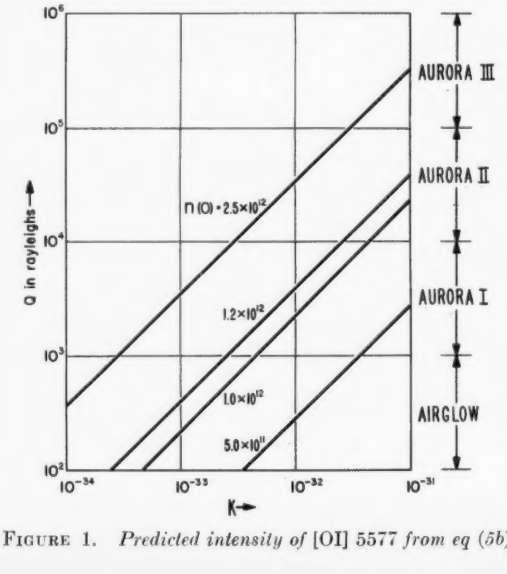


FIGURE 1. *Predicted intensity of* [OI] 5577 *from eq (5b).*

leighs [4] as a function of K for different values of n (0), boxing in the probable range of these parameters. It is seen that already the minimum probable values give a very bright airglow. Either the yield of excited oxygen atoms, (3), has been overestimated, or the importance of collisional deactivation has been underestimated. If neither is the case, it is safe to state that within the accuracy of these data, the Chapman reaction can not only explain the observed airglow intensities, but it actually predicts a faint aurora (1100 rayleighs). However, in addition to explaining the intensity level, a successful theory also has to account for the time variations of the intensity. On a purely static picture, it is not possible to understand why the intensity maximum does not come in the beginning of the night. Since high wind speeds of the emitting cells have been deduced the dynamics of the airglow emission into the photochemical theory will now be introduced to explore the possibility of nocturnal time variations as observed.

3. Meterological Approach

As long as airglow is considered as a steady, static, more or less homogeneous emission from a layer at about 100 km, it is natural to try to explain it in terms of a pure photochemical process solely governed by the kinetic properties of the atmosphere. With the introduction of a dynamical picture in which the airglow emission is concentrated in cells that move relative to the earth's surface, one is led to ask whether the very dynamics of the emissions have anything to do with the excitation of the phenomenon.

At least two avenues lead from here. First, one may ask whether the dynamics of the upper atmosphere can modify the emissions caused by the photochemical reactions to a sufficient degree to account for the observed morphology of the airglow. This may be termed the meteorological approach. Second, one may go a step further and inquire into the possibility of using the dynamics also for the excitation of the oxygen atoms. In this case, the energy of the environmental effects is used as in the next section.

In light of the increased interest of recent years in the study of solar-weather relationships and in the possibility of a coupling between the troposphere and high atmosphere (including the lower part of the ionosphere), it seems appropriate to consider in some detail the meteorological approach to the airglow problem.

The existence of cells in the airglow emission is probable, and from a meteorological point of view, they may be considered as "clouds." The very concept of clouds is synonomous with instability, and the first question is whether one would expect instability in the high atmosphere similar to that in the troposphere. Fleagle [14] has recently shown that disturbances tend to grow rapidly in the mesodecline, at least in polar regions. Also, it should be remembered that the outstanding feature of the troposphere in this connection is the lapse-rate—the decreasing temperature with height and the resulting small Richardson number (Richardson [15]) which is the criterion that turbulence in a gravitational field will increase.[5] (In contrast, the high Richardson number in the stratosphere points to stable conditions.) The Richardson number is given as:

$$R_i=\frac{g\left(\frac{dT}{dz}+\Gamma g\right)}{\left(\frac{dv_H}{dz}\right)^2 T} \tag{9}$$

where Γ is the adiabatic lapse rate $\left(\equiv\frac{1}{c_p}\right)$,

c_p=specific heat at constant pressure,
g=acceleration of gravity,
T=temperature,
v_H=horizontal velocity component of the wind, and
z is in the vertical direction.

As an example, the Richardson number for the lower troposphere, assuming the following characteristics:

$$c_p=1.003\ \text{j/g°K},$$

$$T=270°\ \text{K},$$

$$\frac{dT}{dz}=-5\times10^{-7}\ \text{°K/cm}$$

$$\frac{dv_H}{dz}=1.9\times10^{-2}/\text{sec},$$

is

$$R_i(\text{troposphere})=0.98.$$

The only other region in the atmosphere where a decreasing temperature with height is found is in the mesodecline, at the top of which the airglow phenomenon is observed. Data are not so easily available to compute the Richardson number there, but the vertical change in the wind velocity is of the same order of magnitude as in the troposphere and instability would be expected.[5] In other words, convection is likely to occur and it should reach roughly to the region of minimum temperature, the mesopause, and this is roughly where we find the green airglow.

If the emission changes can be correlated with the convective motions, then, on this picture, the airglow cells are to be considered as the top part of the convection elements in the mesodecline. It is *a priori* not possible to predict in any detail how the penetration of air masses into the atomic oxygen concentration peak will affect the green emission, and only the conjecture is made here that this is the reason for the time variation of the airglow intensities.

---

[4] If the brightness, B, is measured in units of 10^6 quanta/cm^2 second steradian, then the brightness in rayleighs (R) is $4\pi B$.

[5] It is generally agreed (Godske, Bergeron, Bjerknes, and Bundgaard [16]) that a Richardson number somewhat less than unity points to instability. But the exact critical value of the Richardson number is not known.

Now, the mesopause is a rather shallow minimum. One even thinks of it as an isothermal region situated at around 85 km. This is below the airglow height and this looks, at first sight, somewhat inconsistent. Two points should, however, here be taken into consideration. First, as Nicolet [10] points out, the height of the mesopause is not known to the nearest kilometer. Nicolet gives as the best estimate 85 ± 10 km. Second, if the region around 90 km is nearly isothermal, one would expect the convection elements to penetrate up through this layer, similarly as one observes eddies in the region just above the tropopause.

The interpretation of our data as wind implies that we have a component of ordered motion, a sort of "jet-stream," near the mesopause. On the other hand, this motion is certainly superposed on turbulence. The Reynold's number is

$$R_e = \frac{L\bar{v}}{\nu}$$

where L is a characteristic length, $\bar{v}$ is a characteristic velocity, and ν is the kinematic viscosity. With $\bar{v}=10^4$ cm/sec and $\nu=4\times10^4$ cm^2/sec, a Reynold's number $R_e=0.25\ L$ (L in cm) is found. It is not obvious which characteristic length to put in, but if the boundary layers of the cells involved in the turbulence have a thickness as small as 1 km, $R_e=2.5\times10^4$ is found. Convection will tend to inhibit the turbulence, but since this is a region to which the convection only just penetrates, the characteristics of turbulence should probably manifest themselves. Evidence of this may be the rapid changes in luminosity often observed in the airglow cells.

On this picture, the maximum airglow intensity during a night may now come at any time, depending on the penetration of the instabilities into the maximum layer of the atomic oxygen distribution. Since the intensity will vary as $[n(0)]^3$ rather great variations may be expected, and due to the nature of turbulence relatively erratic time variations are not excluded. The green airglow is hence to be considered as a manifestation of "the weather in the mesodecline."

Obviously, too little is known about the dynamics of the regions involved to draw any definite conclusions, but this approach may be worth further study as more data on this part of the atmosphere become available.

Attention is called to a study of 5577 by Tohmatsu [17] in which the Chapman reaction is assumed and the localized variations in atomic oxygen concentration are invoked to account for the observed patchiness of the night airglow, as well as the large scale temporal and geographical charges.

4. Environmental Effects

The fact that the 5577 airglow emission occurs in the ionosphere will now be considered. The electron density is small compared to the neutral gas density, but even in such a slightly ionized plasma new effects are likely to occur. This is the more probable since the plasma is pervaded by and moves in the earth's magnetic field.

The ionosphere has a general similarity to a gas discharge tube, and it is reasonable to inquire whether the conditions that lead to discharges in the laboratory are present in the ionosphere and will lead to luminous discharges that can be associated with airglow.[6] It is characteristic of gas tube discharges that they can be explained only on the assumption that the electrons present are able to absorb energy selectively from an electric field between elastic collisions with heavy particles in the plasma. The essential feature is hence the presence of an electric field capable of accelerating electrons until they have enough energy to excite oxygen atoms. This leads to conditions in the plasma radically different from equilibrium and characterized by an electron gas temperature greatly in excess of the heavy particle temperature.

Several attempts have been made to formulate the problem. Chamberlain [18] *postulated* the existence of an electric field and determined the energy distribution curves and drift velocities of the electrons. He was able to show that auroral rays could be explained on such a discharge theory. Alfvén [19] and Wulf [20, 21] also discussed the mechanism by which the electric field could be set up. The problem includes two main facets: (a) The existence of an electric field capable of accelerating electrons, and (b) the heating of the electron gas and the excitation of the oxygen atoms.

(a) *Production of an electric field.*

Since it is implied here that the dynamics of the 100-km region are of prime importance to an understanding of the green emission, those attempts to establish an electric field which involve extraterrestrial streams (see Alfvén [19]) will not be considered. This stage of the discussion is more in line with Wulf [21], inasmuch as the velocity $\mathbf{V}_c$ of the airglow cells as they move across lines of force of the earth's magnetic field $\mathbf{B}$ will be used. In doing so an electric field is induced

$$\mathbf{E}_{\text{ind}} = \mathbf{V}_c \times \mathbf{B} \text{ (emu)} \qquad (10)$$

This field will tend to produce an electrostatic separation of charges until the polarization field $\mathbf{E}_{\text{pol}}$ thus created exactly balances $\mathbf{E}_{\text{ind}}$. Hence, in the frame of reference of the cell, there is no net electric field. Or, expressed in another way, the polarization field gives the charged particles (irrespective of sign) a drift velocity $\mathbf{V}_d$ in the direction of the motion $\mathbf{V}_c$, such that $\mathbf{V}_d = \mathbf{V}_c$. The plasma therefore moves "downstream" as an entity, and

$$\mathbf{V}_d = -\frac{\mathbf{E}\times\mathbf{B}}{B^2}. \qquad (11)$$

[6] It is a pleasure to acknowledge the many stimulating discussions on this topic with Dr. R. N. Thomas who first drew our attention to this similarity.

In the absence of collisions, the electrons do not see any field and a heating of the electron gas is not possible. When the collisions are taken into account, the problem becomes quite difficult, and it is not at all obvious that the electrons can draw energy from the field. Wulf [21] avoided this basic difficulty by letting the circuit be closed through the F_2 layer where the discharge took place.[7] This would be physically understandable, but it is known the airglow emission comes where the wind speed $\mathbf{V}_e$ is measured, and this rules out the original Wulf approach. It may be recalled that Wulf used eq (10) and found that with a value of 0.5 gauss for the earth's magnetic field and a wind speed of 50 m/sec a gradient of

$$\boldsymbol{E}=0.5\times5\times10^3=2.5\times10^3 \text{ emu}$$

$$\text{or } \boldsymbol{E}\approx10^{-5} \text{ volts/cm}$$

would be created. As mentioned, this potential drop is not available for electron acceleration in this case, since an external path over which to discharge is needed.

Another possibility remains. Roach et al. [5] found evidence of a rotational velocity of the order of 400 m/sec in the airglow cells. If this can lead to polarization effects in the cells, stronger fields (perhaps 10^{-4} to 10^{-3}v/cm) might be expected to be set up between adjacent cells due to the great differential peripheral speeds. But the whole question is for the time being wide open, and no conclusion should be drawn before this problem has been thoroughly analyzed. The best for the moment is to assume the existence of the field and see what the results might be.

(b) *Acceleration of electrons and excitation of oxygen.*

If the same point of view as Chamberlain's [18] is taken, the question may be asked how a postulated electric field will affect the electron energy distribution. According to Chamberlain a gradient of order of magnitude between 10^{-5} and 10^{-4}v/cm is sufficient to account for the acceleration of electrons to an energy capable of exciting the oxygen atoms in auroral rays. Now, it should be remembered that this conclusion is reached on the basis of deriving the electron energy distribution by a simplified theory (Smit [22]). The complete treatment of the problem from the appropriate Boltzmann equation remains to be tackled.

The intensity of the 5577 emission is, according to Chamberlain, given by

$$Q(5577)=\frac{A_{32}}{A_{32}+A_{31}}\,n(\mathrm{OI})n_eS_{13}\frac{H}{10^6} \qquad (12)$$

where collisional deactivation has been neglected, and Q is in rayleighs.

Using a thickness of the emission layer equal to a scale height (H=6.5 km), the values given in the right side of figure 2 ($n_e=10^2$ to 10^5) are obtained; and with H=100 m (to compare with Chamberlain's auroral conditions) the values in the left side of figure 2 (for $n_e=10^4$ to 10^8). An energy ϵ ($\equiv E\lambda$ where λ is the mean free path) of approximately 10^{-1} v is necessary to explain the airglow intensities. The extremely sharp dependence of the emission on this parameter (ϵ) may have an important bearing on the intensity variations in time and space of the airglow pattern, since inhomogeneities in E are likely to occur.

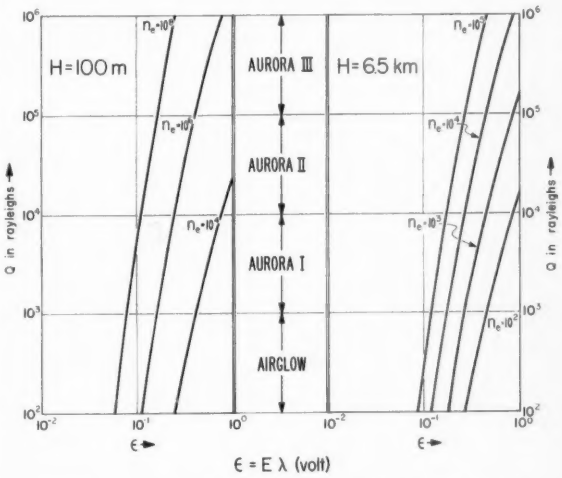


FIGURE 2. *Predicted intensity of* [OI] 5577 *from eq (12).*

The airglow is a normal feature of the "quiet" atmosphere in the 100-km region, and it would here seem reasonable to compare it with other quiet-condition phenomena of electromagnetic nature. The daily variations in the geomagnetic field are such a phenomenon. The variations can be explained in terms of a current system in the upper atmosphere, and rocket observations (Cahill [23]) show the existence of currents at 95- to 110-km altitude. The question may therefore be asked whether electrons could be sufficiently accelerated in the electric field associated with this current system. On the night side of the earth where the airglow is observed, the current between equator and 40° north is approximately 40,000 amp (Chapman and Bartels [24]). The width of the current sheet is hence 40×110 km ≈4,000 km. The thickness l of the sheet is not too well known so the current density j is written as

$$j=\frac{I}{A}$$

[7] At that time, it was believed that the green airglow originated at a height of approximately 200 km.

where the area $A=l\times4\times10^8\text{cm}^2$ (l in cm). Using Ohm's law,

$$j=\sigma\mathbf{E} \tag{13}$$

the corresponding electric field **E** is found provided the anisotropic electrical conductivity σ is known. On magnetically disturbed days additional intense current systems at high latitudes can be estimated Matsushita [25]). Under quiet conditions (Matsuchita [26]) the electric field is estimated to be approximately 3×10^3 emu$=3\times10^{-5}$ v/cm. This is on the low side of what figure 2 indicates is necessary to produce the observed emission. On the other hand, the value 3×10^{-5} v/cm is clearly a lower limit, since it is deduced using quiet geomagnetic conditions. The existence of magnetic disturbances means a greater electric field, and on the discharge theory this may be sufficient for the necessary heating of the electron gas.

It is hardly necessary to stress that this *per se* is no theory for the airglow, but the calculations indicate that it may be worth looking for an explanation of the λ5577 emission along these lines.

5. Summary and Conclusions

The preceding paragraphs have examined existing theories and hypotheses of the green oxygen emission and presented promising alternatives. The sparsity of data available for some of the coefficients and parameters involved makes numerical calculations difficult, but some general conclusions can be drawn.

It is believed that "a modified photochemical theory" is capable of explaining the observed 5577 emission. The complex dynamics of the 90- to 100-km region are not well known, but from what can reasonably be deduced, the "weather conditions" in the mesodecline are compatible with the observed variations of the 5577 emission in space and time.

Similarly, it is possible that the emission can be understood on the basis of "the environmental approach." In particular, the steepness of the curves in figure 2 indicates that small changes in electron density and/or the electric field (which probably occur in the low E layer at night) will lead to rapid changes in brightness, and may explain space and time variations of the emission.

From the present study, *one* mechanism cannot be pointed to as the only possibility. This is perhaps not a very happy situation, but it may be that Nature is using two different mechanisms to generate the observed emissions.

The research reported in this paper has in part been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)–2140 to the High Altitude Observatory.

6. References

[1] J. P. Heppner, J. D. Stolarik, and L. H. Meredith, Rocket measurements at WSPG, New Mexico, July 5, 1956, Trans. Am. Geophys. Union **38,** 394 (1957).
[2] F. E. Roach, L. R. Megill, M. H. Rees, and E. Marovich, The height of nightglow 5577, J. Atmospheric and Terrest. Phys. **12,** 171 (1958).
[3] R. Tousey, Rocket Measurements of the night airglow, Ann. géophys. **14,** 186 (1958).
[4] F. E. Roach, E. Tandberg-Hanssen, and L. R. Megill, The characteristic size of airglow cells, J. Atmospheric and Terrest. Phys. **13,** 113 (1958).
[5] F. E. Roach, E. Tandberg-Hanssen, and L. R. Megill, Movements of airglow cells, J. Atmospheric and Terrest. Phys. **13,** 122 (1958).
[6] M. J. Seaton, Theories of the airglow spectrum, Ann. geophys. **11,** 232 (1955).
[7] S. Chapman, Some phenomena of the upper atmosphere, Proc. Roy. Soc. (London) **A 132,** 353 (1931).
[8] D. R. Bates, The physics of the upper atmosphere, The Earth as a Planet, ed. G. P. Kuiper, ch. 12 (University of Chicago Press, 1954).
[9] D. R. Bates and H. S. W. Massey, The basic reactions in the upper atmosphere I, Proc. Roy. Soc. (London) **A 187,** 261, (1946).
[10] M. Nicolet, La thermosphere, Ann. geophys. **15,** 1 (1959).
[11] R. Garstang, Energy levels and transition probabilities in p^2 and p^4 configurations, Monthly Notices Roy. Astron. Soc. **111,** 115 (1951).
[12] R. A. Minzner, W. S. Ripley, and T. P. Condron, U. S. extension to the ICAO standard atmosphere (1958).
[13] M. J. Seaton, Excitation processes in the aurora and airglow, J. Atmospheric and Terrest. Phys. **4,** 295 (1954).
[14] R. G. Fleagle, Inferences concerning the dynamics of the mesosphere, J. Geophys. Research **63,** 137 (1958).
[15] L. F. Richardson, The supply of energy from and to atmospheric eddies, Proc. Roy. Soc. (London) **A 97,** 354 (1920).
[16] C. L. Godske, T. Bergeron, J. Bjerknes, and R. C. Bundgaard, Dynamic meteorological and weather forecasting, p. 459, (Am. Meteorol. Soc., Boston, Mass., 1957).
[17] T. Tohmatsu, The dynamical effect on the intensity of nightglow, Rept. Ionosphere Research Japan **12,** 253 (1958).
[18] J. W. Chamberlain, Discharge theory of auroral rays, The airglow and the aurorae, ed. E. B. Armstrong and A. Dalgarno (Pergamon Press, 1955).
[19] H. Alfven, Cosmical electrodynamics (Oxford, Clarendon Press, 1950.)
[20] O. R. Wulf, On the relation between geomagnetism and the circulatory motions of the air in the atmosphere, Terrestrial Magnetism and Atmospheric Elec. **50,** 185 (1945).
[21] O. R. Wulf, On the production of glow discharges in the ionosphere by winds, J. Geophys. Research **58,** 531 (1953).
[22] J. A. Smit, Berechnung der geschwindigkeits-verteilung der elektronen bei gasentladungen in helium, Physica **3,** 543 (1936).
[23] L. J. Cahill, Jr., Magnetic exploration of the upper atmosphere, Thesis, State University of Iowa (1959).
[24] S. Chapman and J. Bartels, Geomagnetism, Vol. I, Geomagnetic and related phenomena (Oxford University Press, 1951).
[25] S. Matsushita, Ionospheric variations associated with geomagnetic disturbances, J. Geomag. Geoelectr. **5,** 109 (1953).
[26] S. Matsushita, On possible rocket observations of electric fields, Solar Research Memo No. 125, High Altitude Observatory, Univ. of Colorado (1959).

Boulder, Colo. (Paper 63D3–29)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 3, November–December 1959

A Method for Measuring Local Electron Density from an Artificial Satellite

L. R. O. Storey

(June 23, 1959)

A method is proposed for measuring the electron density at known points in the outer ionosphere, by the use of vlf receiving equipment in an artificial satellite, in conjunction with a vlf transmitter on the ground. The transmitter would radiate continuous waves, which would be propagated through the ionosphere in the "whistler" mode. The basis of the method is a measurement of the local wave admittance of the medium, by comparison of the signals received on an electric dipole and on a loop.

A further proposal is made for an integrated vlf satellite experiment, in which several different types of observation would be made simultaneously.

1. Introduction

Until recently, most of our knowledge of the ionosphere had been obtained by indirect methods, such as vertical-incidence soundings with radio waves. Now, however, the ionosphere can be investigated directly, with equipment borne aloft in rockets or satellites. These two vehicles lend themselves to making complementary studies: a rocket, which goes straight up and then falls back to earth, is better for studies of the variations of the properties of the ionosphere with height at a particular time and place, while a satellite, which may stay in orbit for a long period, is better for synoptic studies of slow variations with time and latitude at particular heights. Rocket and satellite experiments are especially valuable as a source of information about the outer ionosphere, above the peak of the $F2$ layer, which is beyond the reach of conventional vertical soundings.

In studying the ionosphere, the number density of free electrons is a quantity of major interest. Already there are several ways in which the influence of the ionospheric electrons has been observed on transmissions from the artificial satellites; examples are the Faraday rotation of the plane of polarization, and the difference in refraction at disparate frequencies. Such observations, however, all measure integrated effects of the distribution of electrons along the entire path from the satellite to the observer, rather than their density at any one point. But the special advantage of a satellite is that it provides the opportunity for making local measurements at known points in the ionosphere, and so the question arises of how to equip it to measure local electron density. This paper sets out a proposal for one new method. Elsewhere, other methods have been suggested (Hoffmann [1]).[1] Since few have been put to practical test, it is too early to assess their relative merits.

[1] Figures in brackets indicate the literature references at the end of this paper.

The basis of the proposed method is the measurement of the local *wave admittance* of the ionospheric medium. To understand this concept, consider first a linearly-polarized plane wave in free space. The electric ($\underline{E}$) and magnetic ($\underline{H}$) vectors, which comprise the field of the wave, lie at right-angles to each other in the plane of the wavefront (fig. 1). At any fixed point, the variations of the two vectors are always in phase, so that their ratio E/H is a constant. This constant, which has the dimensions of a resistance, is called the *wave impedance of free space* (R_o); its value is about 377 ohms. Its physical significance may be appreciated from the fact that if the wave were to impinge directly upon a uniform sheet of lossy material, that had this resistance per unit square area of surface, then it would be absorbed without reflection. Thus such a sheet would be "matched" to free space.[2]

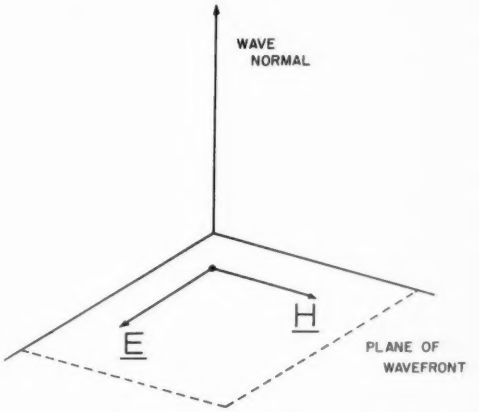


FIGURE 1. *A linearly-polarized plane wave in free space.*

[2] Strictly, for the absorption to be complete the resistive sheet would have to be backed by a perfectly reflecting sheet, at a distance of a quarter-wavelength.

Within a uniform dielectric medium, the ratio of E to H is different from its value in free space. It is given by

$$E/H=R=R_0/\mu \tag{1}$$

where μ is the refractive index. In words, the wave impedance of the medium is inversely proportional to its refractive index. Henceforth, therefore, it will be more convenient to talk in terms of the wave admittance, which is the reciprocal of the impedance, and so is directly proportional to the refractive index:

$$H/E=A=\mu A_0. \tag{2}$$

Evidently, if the wave admittance of the medium could be measured in some way, this measurement would yield the value of the local refractive index.

In applying this result to the measurement of local electron density in the ionosphere, the following experimental arrangement is envisaged: There is to be a transmitter on the ground, radiating continuous waves of constant amplitude. The waves travel upward into the ionosphere, where an orbiting satellite is equipped to receive them. The satellite receives the waves on two separate antennas, a loop for the magnetic field and a dipole for the electric field. Information about the strength of the two fields is telemetered back to ground. At the ground station, the ratio of the two field strengths is computed, giving the wave admittance of the ionosphere in the immediate neighborhood of the satellite. By comparing the observed wave admittance with the value to be expected for free space, the local refractive index is obtained. Now the refractive index is related to the electron density by the expressions of the magneto-ionic theory. Hence, given the local refractive index, it is possible to work backwards and arrive finally at the value of the local electron density. These are the essential features of the proposed experiment, though there are many problems of detail.

The first problem to consider is the choice of the frequency of the transmitted waves. In principle, the experiment would work at any frequency. But if measurements are to be made above the maximum of the $F2$ layer, then the choice is restricted to those frequencies that are capable of passing through this layer: either the frequency must exceed the "critical frequency" of the layer, which varies roughly in the range 5 to 15 Mc, or it must be less than the gyrofrequency (about 1.5 Mc), in which case the waves can penetrate the layer in the "whistler" (ordinary) mode. It is preferable to choose low frequencies and the "whistler" mode, on the following two counts:

First, at low frequencies only this mode is propagated, whereas at high frequencies both the ordinary and extraordinary mode would be propagated; their mutual interference would produce a complicated pattern of field in the ionosphere.

Second, at low frequencies the waves are affected very strongly by the ionospheric electrons, so that the refractive index differs greatly from unity; high frequencies are affected much less. To illustrate this point, suppose that at some level above the $F2$ layer the electron density is such as corresponds to a plasma frequency of 1.5 Mc (see sec. 2.1.). For the transmitted frequencies, take 15 kc as a typical low value and 15 Mc as a typical high value. The refractive indices for these two frequencies are then as given in table 1 below:

TABLE 1. *Refractive indices of the ionosphere, for a plasma frequency of 1.5 Mc*

Wave frequency	Refractive index
15 kc	12.5 (whistler mode).
15 Mc	0.99 (ignoring magnetic field).

In this example, the change of the refractive index from unity is more than a thousand times greater at the lower frequency. There can be no doubt, therefore, that low frequencies are preferable. However, since the "whistler" mode requires the presence of a magnetic field to support its propagation, the use of this mode does confine the experiment to the domain of the earth's magnetic field.

Having chosen the frequency and mode of propagation of the transmitted waves, the remaining problems arise from the fact that, at low frequencies, the ionosphere does not behave at all like a simple dielectric, so that the whistler mode has quite a complicated field structure. The outstanding questions are: (a) What is the structure of the wave field in the whistler mode, and how is the wave admittance defined for this mode? (b) How should the satellite be equipped to explore the structure of the field, and measure the wave admittance? (c) What sort of records would be obtained at the ground? (d) How should the records be interpreted? These questions are discussed successively in sections 2 to 5 of this paper. Section 6 describes a proposal for an integrated vlf satellite experiment, in which several different types of observation would be made simultaneously, so as to obtain the fullest information possible from the vlf transmissions; it includes an independent method for obtaining another measure of the local electron density. Finally, section 7 reviews the proposals and examines some possible sources of error.

2. Structure of Waves in the Whistler Mode

2.1. Introduction

In the proposed experiment, a transmitter on the ground radiates continuous waves at a very low fixed frequency. Some of these waves reach the lower boundary of the ionosphere, where their energy is divided between the two magneto-ionic modes. The ordinary mode is reflected, with some absorption. The extraordinary mode also is partly reflected and partly absorbed, but in addition it experiences a mode coupling effect around the ordinary reflection level. The coupling is brought

about by collisions, and by the rapid variation of the electron density with height in the ionosphere at this low level. Its effect is to transfer some energy from the extraordinary mode through to that branch of the ordinary that is propagated freely at large electron densities; this branch is the mode of propagation of "whistlers." Judging from measurements on the "whistlers" (Storey [2]), energy is coupled into the ionosphere over an area around the transmitter that is about 1,000 km in radius. Thus, there is a vlf field in the ionosphere above this whole area, due to waves that have penetrated the lower boundary and traveled upwards in the "whistler" mode, together perhaps with some downcoming waves, that would arise from partial reflections at any steep gradients of electron density. In the upper atmosphere, where the medium is probably fairly uniform, there is likely to be just a single upgoing plane wave. Moreover, owing to the strong refraction at the base of the ionosphere, the wave is likely to be traveling more or less radially outwards from the earth (Storey [2]). The structure of the field of this wave must be considered now.

In describing the field structure, the rectangular coordinate system that is shown in figure 2 will be adopted. The coordinate axes are labeled 1, 2, and 3, and in that order form a right-handed set: axis 3 lies along the direction of propagation, that is to say, the wave normal; axis 1 is directed so that the geomagnetic field vector $\underline{\mathrm{H}}_0$ lies in the 1–3 plane, and has a positive 1-component: axes 1 and 2 define the plane of the wavefront. The various components of the vectors $\underline{\mathrm{H}}$ and $\underline{\mathrm{E}}$ in the field of the wave will be indicated thus: $\underline{\mathrm{H}}=(H_1, H_2, H_3)$ and $\underline{\mathrm{E}}=(E_1, E_2, E_3)$.

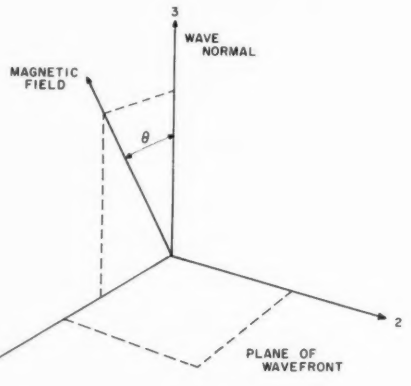


FIGURE 2. *The coordinate system.*

The term "structure" is used here to mean the relative values of the various axial components of the wave field at a point, and how these components vary from point to point and from one instant of time to another. The structure, in this sense, is independent of the amplitude of the wave, because the medium is linear. For determining the electron density, the absolute amplitudes of the field components need not be known, and the only requirement is that they be large enough to overwhelm external atmospheric noise and circuit noise in the receivers. Nevertheless, the absolute amplitudes would be measured in the course of the experiment, and they are of interest for other reasons.

The structure of a plane wave of constant frequency can be described fully by the following quantities:

(a) *Refractive index.* If there is no absorption, all components of the wave vary in time and space as

$$\exp\{j\omega(t-\mu x_3/c)\}$$

where μ is the refractive index, $\omega(=2\pi f)$ is the angular frequency of the wave, t is the time, c is the speed of light in free space, and x_3 is the component of distance along the wave normal.

(b) *Polarization.* The polarization is specified by the ratios of the components of $\underline{\mathrm{H}}$ among themselves (H_1: H_3: H_3), and of the components of $\underline{\mathrm{E}}$ among themselves (E_1: E_2: E_3).

(c) *Wave admittance.* The wave admittance is the ratio of one component of $\underline{\mathrm{H}}$ to one of $\underline{\mathrm{E}}$. Here use will be made of the admittance looking in the direction of propagation, which is defined as H_2/E_1.

These quantities are functions of the frequency of the wave, its direction of propagation relative to the earth's magnetic field, and the local properties of the medium. Expressions for their values are provided by the magneto-ionic theory; the expressions quoted here are taken, with some change of notation, from the work of Booker [3]. Rationalized units are used, with the following definitions:

Constants:

e, charge of electron (a negative quantity)
m, mass of electron
ϵ_0, electric permittivity of free space
μ_0, magnetic permittivity of free space.

Variables:

f, frequency of wave
θ, angle between wave normal and earth's magnetic field
N, number density of electrons
H_0, strength of earth's magnetic field.

Characteristic frequencies:

$$\omega_p^2=(2\pi f_p)^2=\left(\frac{e^2}{\epsilon_0 m}\right)N$$

f_p is the electron plasma frequency.

$$\omega_H=2\pi f_H=-\left(\frac{\mu_0 e}{m}\right)H_0$$

f_H is the electron gyrofrequency.

Dimensionless parameters:

$$X=(\omega_p/\omega)^2=(f_p/f)^2$$
$$Y=\omega_H/\omega=f_H/f$$
$$Y_L=Y\cos\theta$$
$$Y_T=Y\sin\theta.$$

With the aid of this notation, the refractive index, polarization, and wave admittance are discussed in sections 2.2, 2.3, and 2.4, respectively. The influence of collisions is ignored, because they are infrequent in the outer ionosphere.

To illustrate the results of the theory, and to examine the adequacy of any approximations, numerical calculations will be made for the following particular set of conditions:

$$f=15.5\text{ kc} \qquad f_p=1.55\text{ Mc}$$
$$\theta=160° \qquad f_H=1.06\text{ Mc}.$$

The parameters of the magneto-ionic theory then have the values

$$X=10^4 \qquad Y_L=-64.5$$
$$Y=68.7 \qquad Y_T=23.5.$$

Such conditions might apply at a height of about 1000 km above the vlf radio transmitter NSS (geomagnetic latitude 50° N). The wave is assumed to be traveling straight out from the earth in the plane of the magnetic meridian, so that axis 3 of the coordinate system is vertical, while the other two axes are horizontal; axis 1 points towards magnetic north, and axis 2 towards magnetic east. Note that in this instance the angle θ is greater than 90°, since the earth's magnetic field is directed downwards in the northern hemisphere.

2.2. Refractive Index

When collisions are ignored, the general expression for the refractive index becomes

$$\mu^2=1-\frac{2X(1-X)}{2(1-X)-Y_T^2\pm\sqrt{Y_T^4+4Y_L^2(1-X)^2}} \tag{3}$$

in which the positive sign corresponds to the ordinary mode, and the negative to the extraordinary. At wave frequencies that are low compared to both the local plasma frequency and gyrofrequency, the refractive index for the "whistler" (ordinary) mode is usually given quite accurately by the quasi-longitudinal (QL) approximation:

$$\mu^2\cong 1+\frac{X}{|Y_L|-1}. \tag{4}$$

The condition for the approximation to hold is that

$$|(X-1)Y_L|>>\tfrac{1}{2}Y_T^2. \tag{5}$$

This equality is well satisfied in the example chosen, where the error of the QL approximation to the refractive index is less than 0.1 percent.

The further approximation

$$\mu^2\cong\frac{X}{|Y_L|}=\frac{f_p^2}{f f_H|\cos\theta|} \tag{6}$$

holds if $X>>|Y_L|>>1$. In the example, this approximation gives μ with an error of about 1 percent.

2.3. Polarization

Consider first the polarization in the plane of the wavefront. The ratios of the transverse components of $\underline{H}$ and $\underline{E}$ among themselves are given by the "polarization ratio"

$$R=-H_2/H_1=E_1/E_2 \tag{7}$$
$$=-jk\{1\pm\sqrt{1+k^{-2}}\} \tag{8}$$

where

$$k=\frac{g}{1-X} \tag{9}$$

and

$$g=\frac{1}{2}\frac{Y_T^2}{Y_L}=\frac{1}{2}Y\frac{\sin^2\theta}{\cos\theta} \tag{10}$$

and where the positive and negative signs correspond to the ordinary and extraordinary modes respectively.

Now the "whistler" mode is the branch of the ordinary mode on which X is greater than unity, so that the denominator in (9) is negative; hence k is negative or positive according to whether θ is less or greater than 90°. Also, at vlf the quasi-longitudinal condition (5) is satisfied, and this condition implies that $|k|\ll 1$. With these restrictions on the value of k, the polarization ratio for the whistler mode is approximately

$$R\cong+j\{1+|k|\} \tag{11}$$

where the sign is positive or negative according as θ is less or greater than 90°, that is, according as the direction of propagation makes an acute or obtuse angle with the direction of the earth's magnetic field.

In general, the polarization is elliptical, with the axes of the ellipse parallel to axes 1 and 2 of the coordinate system; this conclusion follows from the fact that R is purely imaginary, so that H_1 and H_2 are in phase quadrature. The vector sum traces out an ellipse with its minor axis parallel to axis 1. The sum of E_1 and E_2 traces out an ellipse of the same shape, but in this case the minor axis is parallel to axis 2.

The axial ratios of these ellipses are very close to unity; in the example, they differ from unity by less than 0.1 percent. So the polarization is almost circular, as it always is under QL conditions.

The sense of rotation around the circle is specified by the sign of the imaginary value of R, a positive sign meaning right-handed rotation. For the whistler mode, the rotation is right-handed when θ is less than 90°, and left-handed when θ is more than 90°. Of course, the senses of rotation for the $\underline{H}$ and $\underline{E}$ vectors are the same.

Now consider the longitudinal field components, the components along axis 3. The $\underline{H}$ vector has no such component; indeed, radio waves in a plasma have the general property that their magnetic field is always wholly transverse to the direction of propagation. Their electric field, however, in general does have a longitudinal component, which is given by

$$E_3/E_2 = j\frac{(\mu^2-1)Y_T}{1-X} \tag{12}$$

$$\cong -j\left(\frac{Y_T}{|Y_L|-1}\right)\left(\frac{X}{X-1}\right). \tag{13}$$

The approximation holds under QL conditions. Furthermore, if $X \gg 1$ and $|Y_L| \gg 1$,

$$E_3/E_2 \cong -j\frac{Y_T}{|Y_L|} = -j|\tan\theta|. \tag{14}$$

To the same approximation

$$E_3/E_1 \cong -\tan\theta. \tag{15}$$

These expressions mean that E_3 is in phase with E_1, and of such a magnitude and sign that the sum of these two components is roughly perpendicular to the geomagnetic field; in the numerical example, the angle between the sum and the perpendicular is about 3/4°. Now E_2 also lies in the plane perpendicular to the field, so that the total $\underline{E}$ vector traces out an ellipse in this plane, with the minor axis parallel to axis 2 of the coordinate system. The explanation of this result is that the electrons in the ionosphere are able to move much more rapidly in the direction of the geomagnetic field than at right angles to it, and at low frequencies they have sufficient time during a period of oscillation to move and neutralize any component of the electric field in this direction. Thus the result is not true in general, but only at low frequencies.

2.4. Wave Admittance

The admittance looking in the direction of propagation is defined as

$$A = H_2/E_1 = -H_1/E_2. \tag{16}$$

It is given by

$$A = \mu A_0 \tag{17}$$

where μ is the refractive index, given by (3), and A_0 is the admittance of free space

$$A_0 = (\epsilon_0/\mu_0)^{\frac{1}{2}} = (377 \text{ ohms})^{-1}. \tag{18}$$

These expressions show that the field components H_2 and E_1 are in phase. From this and previous results it follows that, in the plane of the wavefront, the rotating magnetic vector leads the projection of the electric vector by 90° when there is propagation along the earth's magnetic field ($\theta < 90°$), and lags it by 90° when the propagation is against the field ($\theta > 90°$).

2.5. Summary of Approximations

To summarize, the structure of waves in the whistler mode can be represented quite well by the following approximate expressions:

For the refractive index,

$$\mu^2 = 1 + \frac{X}{|Y_L|-1}. \tag{19}$$

For the polarization,

$$H_1 = \pm jH_2 \tag{20}$$

$$E_1 = \pm jE_2 \tag{21}$$

$$E_3 = -E_1 \tan\theta \tag{22}$$

where the sign in (20) and (21) is positive or negative according as θ is less or greater than 90°. The assumptions now are that the $\underline{H}$ vector traces out an exact circle in the plane of the wavefront, and the $\underline{E}$ vector an ellipse in the plane exactly perpendicular to the earth's magnetic field, such that the projection of the ellipse onto the plane of the wavefront is a circle also. This state of polarization is illustrated in figure 3; here θ has been taken as acute, for the sake of clarity, whereas this angle is obtuse in the worked example.

Finally, for the wave admittance,

$$A = H_2/E_1 = \mu A_0 \tag{23}$$

with μ given by (19).

These approximations are used throughout the rest of this paper.

3. Details of the Satellite

3.1. General Description

The satellite would be equipped with two separate antennas for receiving the vlf waves. One would be a loop, responding to the magnetic field of the

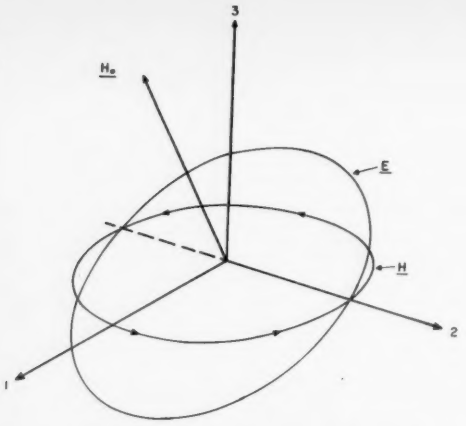


FIGURE 3. *The polarization of waves in the whistler mode.*

wave, while the other would be a dipole, for the electric field. The two antennas would be mounted coaxially, that is to say, the plane of the loop would be perpendicular to the axis of the dipole. With this arrangement, the satellite would receive components of the magnetic and electric vectors resolved in the same direction, that of the common axis of its two antennas. The mechanical design and layout of the satellite are shown in figure 4.

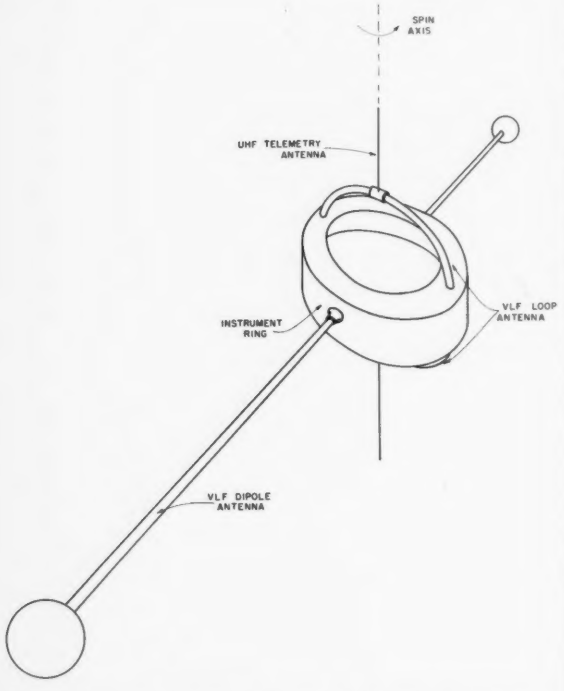


FIGURE 4. *The mechanical design of the satellite (the outer cover is not shown).*

In order to explore the structure of the wave field, the satellite would be set spinning, so that the direction of the antenna axis, and hence the amplitude and phase of the received signals, would change continually. Information is required on the variations of the amplitudes of the signals and of the sign of their correlation. These variations would be easiest to interpret if the spin axis of the satellite was at right angles to the axis of the antennas; then the motion of the antenna axis would be a simple rotation in a plane, and all variations would be periodic. The theory for interpreting these variations, as developed in sections 4 and 5, assumes that this relation between the spin axis and the antenna axis is established exactly.

Details of various aspects of the design of the satellite now follow.

3.2. VLF Antennas

The design of the loop antenna is quite straightforward. At vlf, the most efficient arrangement is an air-cored loop, connected to the receiver through a matching transformer. The basis for optimum design, subject to limitations of weight and space, is provided by the work of Helliwell [4].

The design of the electric dipole is more difficult, since this antenna is in direct contact with the ionospheric plasma, and its properties are affected thereby. The aims of design are to ensure that the effective height of the antenna is the same in the medium as in free space, and that its impedance does not vary excessively as the satellite spins. Consideration of the probable effects of the medium suggest that the best arrangement is one where each arm of the dipole consists of a conducting sphere supported by a long insulating rod; connection is made to the sphere by a thin wire running down the center of the rod. The merits of this arrangement are that the sensitive parts of the antenna are well localized, and are removed as far as possible from the disturbing influence of the body of the satellite. Furthermore, the effective height of this antenna should not be altered by the formation of a "positive-ion sheath" around the satellite (Seddon [5]), though its impedance doubtless would be altered. To minimize the effects of changes of antenna impedance, the input impedance of the receiver should be as high as possible.

3.3. Electronic Equipment

A block diagram of the electronic equipment is given in figure 5. It comprises vlf receivers, detectors, a telemetry system, and a generator for calibration signals.

Two separate vlf receivers are needed, one for each antenna. They should be designed for low noise, stability of phase shift, and wide dynamic range. Their pass-bands should be fairly narrow, to reduce atmospheric noise, but not so narrow that the phases of their outputs fail to reproduce the changes in phase of the received signals that accom-

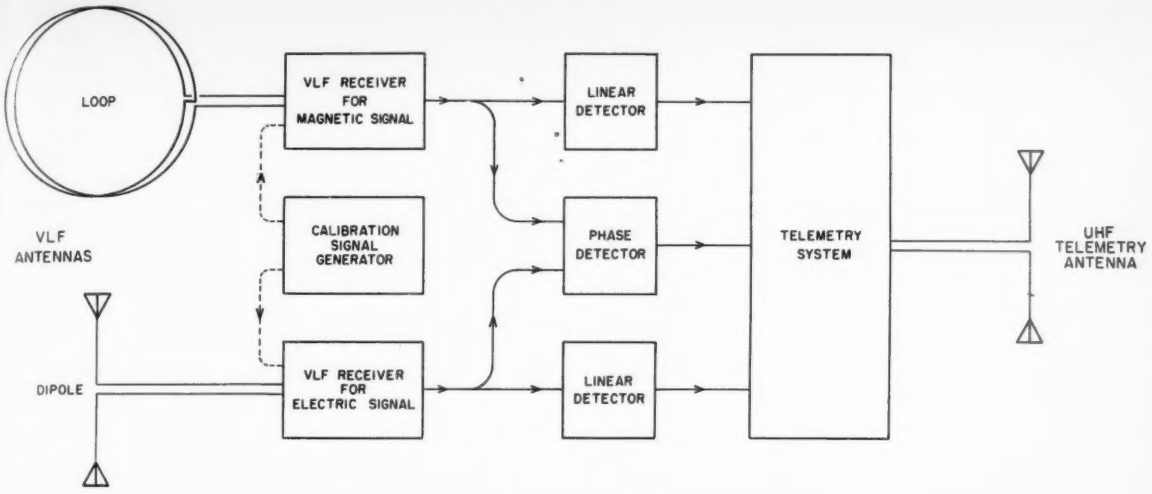


FIGURE 5. *The electronic equipment of the satellite.*

pany the spinning of the satellite. They should incorporate negative feedback to improve both the stability and the dynamic range.

The amplitude of the signal that emerges from each receiver is measured by a detector, the output from which is telemetered to ground. Also, the two signals are compared in a phase detector, and the polarity of the output relayed to ground on a third telemetry channel. The ground station records these three data, which are proportional to the amplitudes of the received signals and the sign of their correlation. The factors of proportionality would have to be determined by calibration of the whole system prior to flight. Also, to confirm that the system is working properly during flight, it should be calibrated automatically from time to time by injecting standard signals at the inputs to the receivers.

3.4. Orbital and Spin Motion

The orbit of the satellite should be eccentric, so that measurements are obtained over a range of height; an apogee height of 1 to 2000 km and a perigee height of about 300 km would be suitable. At this height, the perigee would lie below the maximum of the $F2$ layer from time to time, and then the measurement of electron density could be checked by comparison with ionospheric soundings. There are no special requirements for the inclination, except that it must be large enough to allow the orbit to pass over the transmitter, and that the value 63.4°, at which the line of apsides does not precess, should be avoided.

The necessary rate of spin is determined by the requirement that the satellite, in moving along its orbit, should not experience much change in local electron density during one rotation. If the distribution of electrons in the outer ionosphere is assumed to contain no irregularities smaller in scale than 1 km, and the speed of the satellite is taken to be roughly 5 kmps, then the rate of spin should not be less than 5 rps. If the smallest irregularities actually are larger than assumed here, then proportionately lower rates would be acceptable.

The absolute direction of the spin axis in space should be chosen so that the satellite is oriented favorably for making measurements at the times when it passes over the transmitter. At such times, the spin axis should be more or less horizontal, and should lie roughly in the plane of the orbit. It has to be horizontal so that the spinning motion will bring the antenna axis through the vertical, where the electric dipole is coupled strongly to the field component E_3; for it is on this coupling that the measurement of the angle θ depends (see secs. 4 and 5). The spin axis also has to lie close to the plane of the orbit, to ensure that the elements of the electric dipole do not pass behind the body of the satellite, where they might encounter a wake of disturbed ionization. However, it is doubtful whether such a wake would be formed in the outer ionosphere, where the mean free path of the particles is much larger than the dimensions of the satellite, so this requirement may not be crucial.

3.5. Mechanical Design

Bracewell and Garriott [6] have considered the influence of mechanical design on the free rotation of a satellite. They point out that the characteristics of the free rotation are determined by the relative values of the principal moments of inertia, and that if one moment is much larger than the other two, then the satellite will tend to spin about the axis associated with that moment. Therefore, the mass of the satellite should be arranged so that its moment of inertia is greatest about the desired axis of spin. If a satellite of this type was set spinning initially in a direction slightly incorrect, then its

subsequent motion would include a wobble at twice the spin frequency. This wobble could be damped out by such means as suitably disposed tubes of viscous fluid, whereupon the satellite would be left spinning smoothly about the correct axis.

In the present case, the main aim is to ensure that the spin axis is transverse to the axis of the vlf antenna system. If this were the only aim, and the particular location of the spin axis in the transverse plane could be arbitrary, then it would follow that the satellite should be long and thin, extended along the axis of the vlf antenna system. However, it is better to fix the direction of the spin axis in the frame of the satellite completely, so as to provide a stable direction for the antenna of the telemetry system; to this end the mass should be distributed as in a disk, to make the moment of inertia greatest about the chosen direction.

Once the spinning motion has been established about the desired axis in the frame of the satellite, the absolute direction of this axis in space should remain fixed by gyroscopic action. However, if any force were to exert a systematic couple on the satellite as it moved along its orbit, then the spin axis would precess. Possibly the forces due to atmospheric drag, or to electric currents induced in the body of the satellite as it spins in the earth's magnetic field, could exert such couples. The couple due to atmospheric drag could be eliminated by designing the satellite with reflection symmetry about a plane perpendicular to the spin axis, while induced currents could be prevented by making the body of the satellite in several parts that were insulated from one another.

A mechanical design that embodies these principles is illustrated in figure 4; in this drawing, the protective outer shell of the satellite is supposed to have been removed.

4. Received Signals and Their Variations

4.1. General Behaviour

This section deals with the signals received on the two vlf antennas, and how they vary as the satellite spins. Specifically, expressions are derived for the amplitudes of the two signals and for the sign of their correlation, since this is the information that is relayed from the satellite to the ground.

Now the loop antenna and the electric dipole antenna are set on a common axis (sec. 3), so the signals that they respond to are, respectively, the components of $\underline{H}$ and of $\underline{E}$ along this axis; call these components H_a and E_a. Take the coordinate system of figure 2, and let the instantaneous direction of the antenna axis be specified by its direction cosines a_1, a_2, and a_3 relative to the coordinate axes. Then the received signals are

$$\textit{magnetic signal: } H_a = a_1 H_1 + a_2 H_2 \tag{24}$$

$$\textit{electric signal: } E_a = a_1 E_1 + a_2 E_2 + a_3 E_3. \tag{25}$$

It is convenient to express H_a in terms of H_2 alone, and E_a in terms of E_1, while recalling that the field components H_2 and E_1 are in phase, and that their ratio is the wave admittance (sec. 2.4.). In eliminating the other components, the structure of the field will be assumed to be that given by the approximations of section 2.5. After making the appropriate substitutions, the following expressions are obtained for the received signals:

$$\textit{magnetic signal: } H_a = \{a_2 \pm j a_1\} H_2 \tag{26}$$

$$\textit{electric signal: } E_a = \{a_1 - a_3 \tan\theta \mp j a_2\} E_1. \tag{27}$$

The sign depends on the direction of propagation with respect to the geomagnetic field; the upper sign applies when there is a component of propagation along the field.

As the satellite moves along its orbit, and spins about its axis, the received signals vary in amplitude and phase. The variations due to the orbital motion are slow and systematic, while those due to the spinning motion are relatively rapid and are, for the most part, periodic. In considering the variations over a single rotation period of the satellite, only the effects of spin need to be considered.

One general feature of these variations can be seen immediately; it is that they all must be periodic at twice the spin frequency, since a reversal of the direction of the antenna does not alter the amplitudes of the received signals, nor their relative phase (except possibly by 360°).

To find the exact form of the variations, expressions must be obtained for the direction cosines of the antenna axis as functions of time.

4.2. Direction of the Antenna Axis

The geometry of the situation is represented in figure 6. Figure 6a is a perspective drawing that gives the notation for the angles that specify the directions of the various vectors. In figure 6b the same information is given in the form of a stereographic projection onto the plane of the wavefront. Details of this projection can be found in any standard work on crystallography, such as that of Tutton [7] or Bunn [8]. It maps the surface of a sphere into the interior of a circle on a plane, and has the property that arcs of great circles on the sphere are represented by arc of circles on the plane; those representing great circular arcs on the upper hemisphere are drawn as solid lines, while the projections of arcs on the lower hemisphere are drawn as broken lines. Points in the upper hemisphere are represented by solid dots, and in the lower hemisphere by small open circles. In the present application, the orientations of vectors and planes in space are represented by the projections of their intersections with an imaginary sphere that is concentric with the satellite.

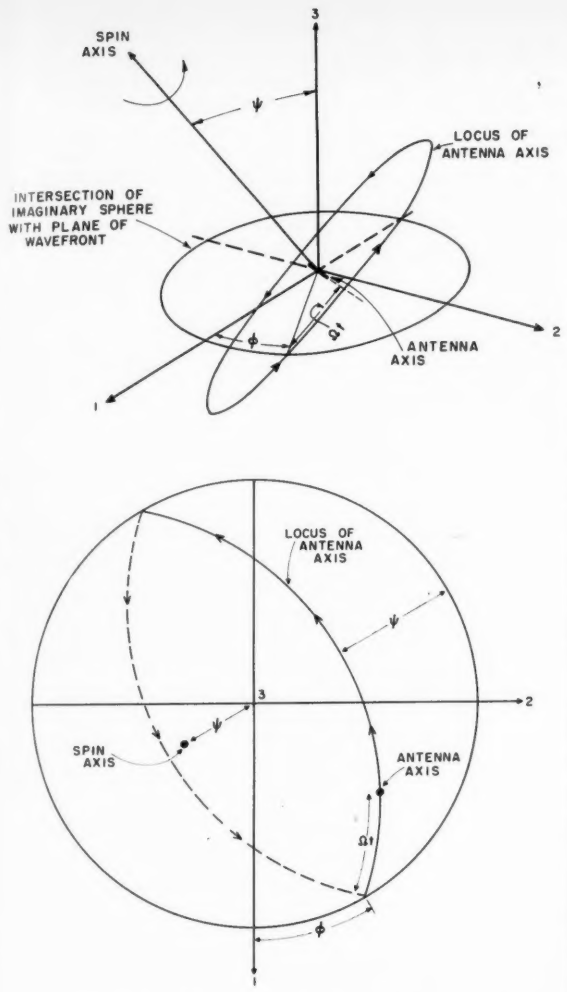


FIGURE 6. *The geometry of the spinning motion.*

(a) In perspective, (b) on a stereographic projection.

The direction of the spin axis of the satellite is taken as that direction about which the rotation is right-handed. It is assumed to be perpendicular to the antenna axis, as specified in section 3. The angle that it makes with axis 3 of the coordinate system is called Ψ, and the angular velocity of the spinning motion is called Ω.

The time t is measured from an instant when the antenna axis lies in the plane of the wavefront (the 1–2 plane). The position of the antenna axis in this plane is then given by the angle ϕ, which is measured from axis 1 to that end of the antenna axis that is going upwards, toward axis 3, at this particular instant. It is reckoned positive if it is measured toward axis 2, and negative otherwise; thus, in figure 6, ϕ is positive.

In these terms, the three direction cosines are given by the expressions

$$a_1=\cos\phi\,\cos\Omega t-\sin\phi\,\cos\Psi\,\sin\Omega t \tag{28}$$

$$a_2=\sin\phi\,\cos\Omega t+\cos\phi\,\cos\Psi\,\sin\Omega t \tag{29}$$

$$a_3=\sin\Psi\,\sin\Omega t. \tag{30}$$

The angles Ψ, ϕ, and θ define completely the relative directions of the spin axis, the wave normal, and the geomagnetic field. They vary progressively as the satellite moves along its orbit, but these variations are slow compared to the spinning motion itself.

4.3. Amplitude of the Magnetic Signal

The real and imaginary parts of the magnetic signal, relative to H_2 as the standard of phase, may be found by substituting the expressions for the direction cosines into (26). Their values, normalized with respect to H_2, are

$$\mathrm{R}\{H_a/H_2\}=a_2=\sin\phi\,\cos\Omega t+\cos\phi\,\cos\Psi\,\sin\Omega t \tag{31}$$

$$\mathrm{I}\{H_a/H_2\}=\pm a_1=\pm(\cos\phi\,\cos\Omega t-\sin\phi\,\cos\Psi\,\sin\Omega t) \tag{32}$$

Thus the square of the normalized peak amplitude of H_a is

$$|H_a/H_2|^2=a^2{}_1+a^2{}_2 \tag{33}$$

$$=1-\sin^2\Psi\,\sin^2\Omega t \tag{34}$$

Note that this expression does not involve either ϕ or θ. It is made up of a constant part, and a part that varies sinusoidally at twice the spin frequency. The normalized peak amplitude, given by the square root of (34), takes its maximum value of unity when the antenna lies in the plane of the wavefront ($\Omega t=0,\ \pi$), and its minimum value of $|\cos\Psi|$ at points halfway in between ($\Omega t=\pi/2,\ 3\pi/2$).

4.4. Amplitude of the Electric Signal

Similarly, the real and imaginary parts of the electric signal E_a, normalized with respect to E_1, are

$$\mathrm{R}\{E_a/E_1\}=a_1-a_3\,\tan\theta$$
$$=\cos\phi\,\cos\Omega t-(\sin\phi\,\cos\Psi+\tan\theta\,\sin\Psi)\,\sin\Omega t \tag{35}$$

$$\mathrm{I}\{E_a/E_1\}=\mp a_2=\mp(\sin\phi\,\cos\Omega t+\cos\phi\,\cos\Psi\,\sin\Omega t). \tag{36}$$

Hence the square of the normalized peak amplitude is

$$|E_a/E_1|^2=(a_1-a_3\,\tan\theta)^2+a^2{}_2 \tag{37}$$

$$=1-(1-\tan^2\theta)\,\sin^2\Psi\,\sin^2\Omega t-2\,\tan\theta\,\sin\Psi\,\sin\Omega t(\cos\phi\,\cos\Omega t-\sin\phi\,\cos\Psi\,\sin\Omega t. \tag{38}$$

This is more complicated than the corresponding expression for the magnetic signal, because the electric field has a longitudinal component. However, it also represents the sum of a constant term, and a sinusoidal term of angular frequency 2Ω. The normalized amplitude of E_a also is equal to unity when $\Omega t=0$ or π, while its value at the points halfway in between is

$$|E_a/E_1|^2=\cos^2\Psi+\tan\theta\sin\phi\sin 2\Psi+\tan^2\theta\sin^2\Psi. \quad (39)$$

In general, however, these are not its maximum and minimum values. The times of stationary amplitude are separated by one quarter of a rotation, and are given by

$$\Omega t=\tfrac{1}{2}\cot^{-1}\{\sec\phi\ (\sin\phi\cos\Psi-\cot 2\theta\sin\Psi)\}. \quad (40)$$

The question of which time corresponds to the maximum may be decided by inspecting the geometry of the situation. In most situations, maximum amplitude occurs near to the time when the antenna axis passes through the plane perpendicular to the geomagnetic field, because this is the plane of the electric vector.

4.5. Sign of the Correlation

Here the term "correlation" is being used rather loosely, to mean the average of the instantaneous product of the signals H_a and E_a, taken over one complete cycle of the radiofrequency. If the two signals were in phase, its value would be $\frac{1}{2}\,|H_a|\,|E_a|$. In general, it is

$$\overline{H_aE_a}=\tfrac{1}{2}[\mathrm{R}\{H_a\}\mathrm{R}\{E_a\}+\mathrm{I}\{H_a\}\mathrm{I}\{E_a\}] \quad (41)$$

$$=\tfrac{1}{2}\,|H_a|\,|E_a|\cos(\alpha_H-\alpha_E) \quad (42)$$

where α_H and α_E are the phases of the magnetic and electric signals, respectively, measured in relation to the phase of H_2 or E_1. From (42), it appears that the correlation is positive when $(\alpha_H-\alpha_E)<90°$, negative when $(\alpha_H-\alpha_E)>90°$, and zero when the two signals are exactly in phase quadrature.

The correlation could be determined by applying the voltages that represent H_a and E_a to the inputs of an electronic multiplier, and then smoothing the output with a time constant that is long compared to the period of the radiofrequency, but short compared to the period of the satellite spin. The proposed satellite would probably use some simpler type of phase detector, rather than a true multiplier. The output from such a detector would not necessarily be proportional in magnitude to the correlation, but it would have the same sign. For the present purpose, only the sign is of interest, so it is satisfactory to discuss the correlation.

The correlation of the received signals is given as a function of the direction cosines by the expression

$$\frac{\overline{H_aE_a}}{\frac{1}{2}|H_2|\quad|E_1|}=-a_2a_3\tan\theta \quad (43)$$

which contains the correlation of H_2 and E_1 as a normalizing factor. From this expression it appears that the sign of the correlation depends jointly on two circumstances: first, whether θ is greater or less than 90°, and second, which of four quadrants of the sphere contains the direction of the antenna axis. These quadrants are bounded by the 1–2 plane and the 1–3 plane (fig. 7a). When the antenna axis lies in either of these planes, passing from one quadrant to another, the correlation is zero. The rules for the signs are given in figure 7b.

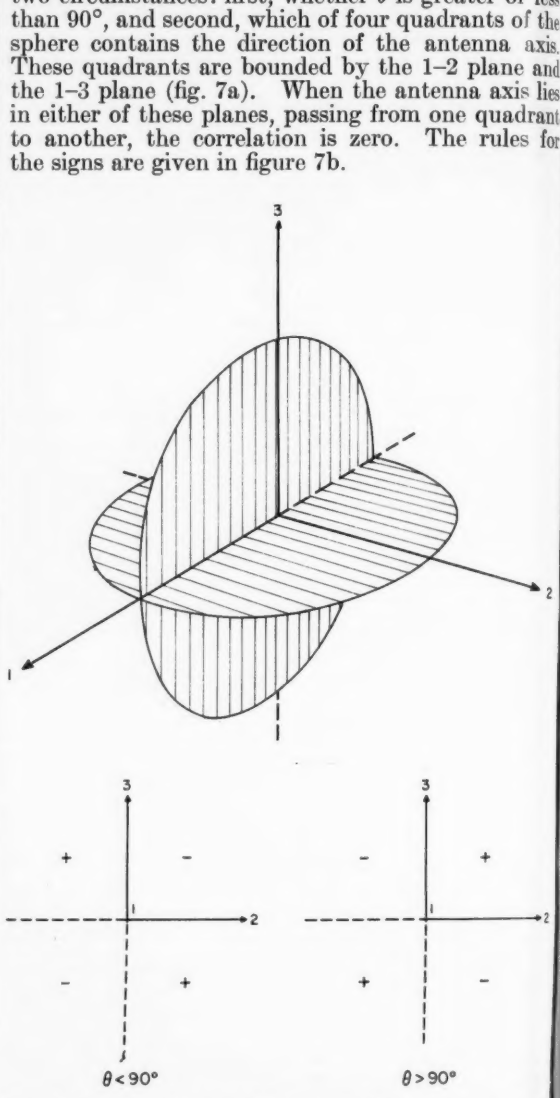


FIGURE 7. *How the sign of the correlation is governed by the direction of the antenna axis.*

(a) The four quadrants of the sphere, within each of which the sign is constant
(b) the rules for the sign.

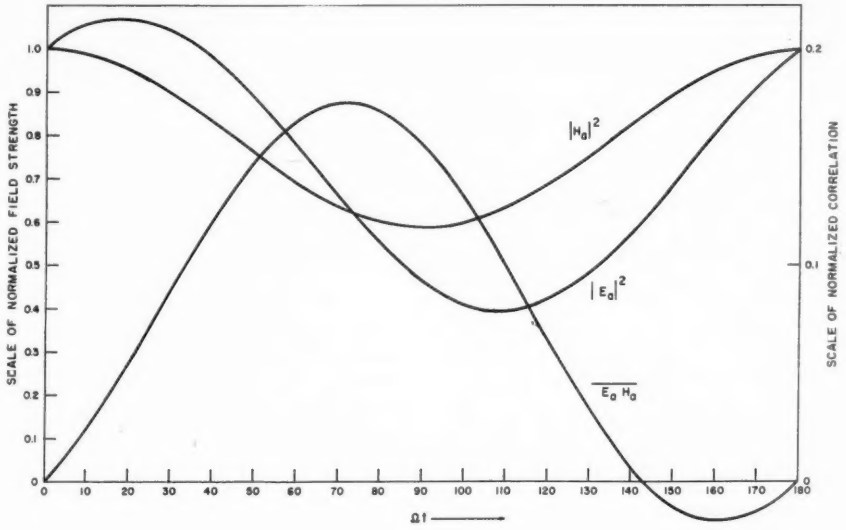
These results can be explained as follows: The departures of the relative phase of H_a and E_a from the quadrature condition are due entirely to the longitudinal component E_3 of the electric field. If this component was absent, as indeed is the case when $\theta=0$, then the total $\underline{\mathrm{E}}$ vector would be wholly transverse and would execute the same motion as the $\underline{\mathrm{H}}$ vector, with a phase lag or lead of 90°; the received signals H_a and E_a, would then be in phase

quadrature for all directions of the antenna axis. When E_3 is finite ($\theta \neq 0$), the signals are exactly in quadrature in just two special situations: the first is the situation where the antenna axis lies in the 1–2 plane ($\Omega t = 0, \pi$), so that the electric dipole is not coupled to E_3: the second is the situation where the antenna axis lies in the 1–3 plane, so that the dipole is not coupled to E_2. Then, since E_3 is in phase with E_1 (see sec. 2.5), the electric signal has the same phase as it would if E_3 were absent. The times of passage of the antenna axis through the 1–3 plane are given by

$$\Omega t = \tan^{-1}\{-\sec\Psi \tan\phi\}. \qquad (44)$$

4.6. An Example of the Variations

The example of figure 8 shows how the properties of the received signals are expected to vary during one half-revolution of the satellite, for a particular orientation of the spin axis. Figure 8a is a graph of the normalized values of the signal strengths $|H_a|^2$ and $|E_a|^2$, and of the correlation $\overline{H_a E_a}$, as observed at the satellite; note that these three quantities all vary sinusoidally. Figure 8b shows how the corresponding records of $|H_a|$, $|E_a|$, and the sign of $\overline{H_a E_a}$ would appear at the ground station.



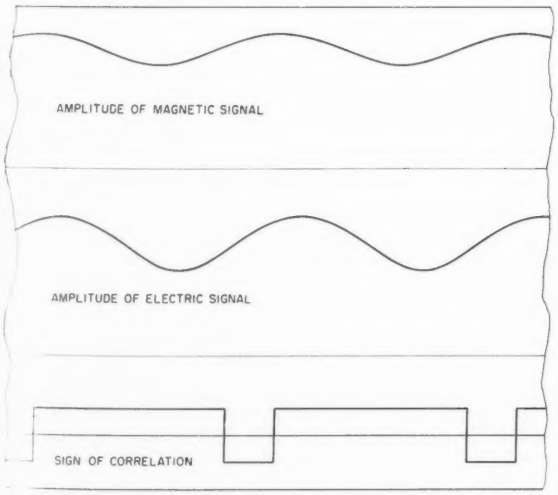


FIGURE 8. *The variations of the properties of the received signals, for the particular case where $\Psi = 40°$, $\phi = 30°$, and $\theta = 160°$.*

(a) The variations of $|H_a|^2$, $|E_a|^2$, and $\overline{E_a H_a}$, (b) the corresponding records.

5. Analysis of the Records

5.1. Wave Admittance, Refractive Index, and Electron Density

Records such as that of figure 8b contain almost all the information that is needed for calculating the electron density of the ionosphere in the neighbourhood of the satellite. The calculation, though, is somewhat indirect.

The quantity that can be obtained most directly from the records is the local wave admittance, A. This, it will be recalled, is the ratio of the field components H_2 and E_1. However, since the polarization in the plane of the wavefront is circular, both for $\underline{\mathrm{H}}$ and for $\underline{\mathrm{E}}$ (sec. 2.3), the wave admittance is given also by the ratio of the electric signal E_a, taken at any instant when the antenna axis passes through this plane. Such instants can be recognized from the fact that the magnetic signal is then at a maximum.

Given the wave admittance, the local refractive index μ can be calculated immediately from (23). Thus

$$\mu = A/A_0. \qquad (45)$$

The relationship between μ and the local electron density N is contained in (19). After rearrangement to express N as a function of the other variables, this equation reads

$$N=\left(\frac{\epsilon_0 m}{e^2}\right)\omega\left\{H_0\left|\frac{\mu_0 e}{m}\cos\theta\right|-\omega\right\}(\mu^2-1). \quad (46)$$

A convenient approximation to this expression may be obtained rearranging (6) similarly; it is

$$N=\frac{\omega}{c^2|e|}H_0|\cos\theta|\mu^2 \quad (47)$$

in which c is the speed of light. Clearly, to calculate the electron density from a given value of μ, it is necessary to know also the local magnetic field strength H_0, and the angle θ between the direction of propagation and the field.

The magnetic field strength cannot be determined from the records. If the position of the satellite is known, the field strength at this point can be estimated roughly from the known distribution of field on the earth's surface, by extrapolating upwards. It would be better, however, to measure the field directly by means of a magnetometer borne in the satellite; this possibility is discussed in section 6.

The angle θ can be obtained from the details of the variations of the recorded data. The argument, which is rather circuitous, depends on the following results from the previous section: (a) The variations of the amplitude of the magnetic signal are governed by Ψ only (34), (b) the times of change of the sign of the correlation are governed by Ψ and ϕ, but are independent of θ (44), (c) the variations of the amplitude of the electric signal depend on all three angles, including θ (38).

Evidently, by analysing these three variations in the given sequence, it is possible first to calculate the angle Ψ, then the angle ϕ, and finally to obtain the angle θ. These calulations are discussed in the rest of this section. The discussion is illustrated by analysing the record of figure 8b.

5.2. The Angle Ψ

Ψ is the angle between the spin axis and the wave normal (sec. 4.2), and its range of possible values is 0 to 180°. The value of Ψ may be obtained from the observed variations of the amplitude of the magnetic signal; in fact, from (34), it is given by the ratio of the extreme values of the amplitude. That is,

$$|\cos\Psi|=\frac{|H_a|\ \min}{|H_a|\ \max}. \quad (48)$$

Hence Ψ is determined, though with ambiguity. For a given value of $|\cos\Psi|$, there are two possible values of Ψ; their sum is 180°.

Thus, in the example of figure 8, the ratio of the minimum and maximum values of $|H_a|$ is 0.766; on this evidence, Ψ is 40° or 140°. The possible locations of the direction of the spin axis are shown on the stereographic projection of figure 9.

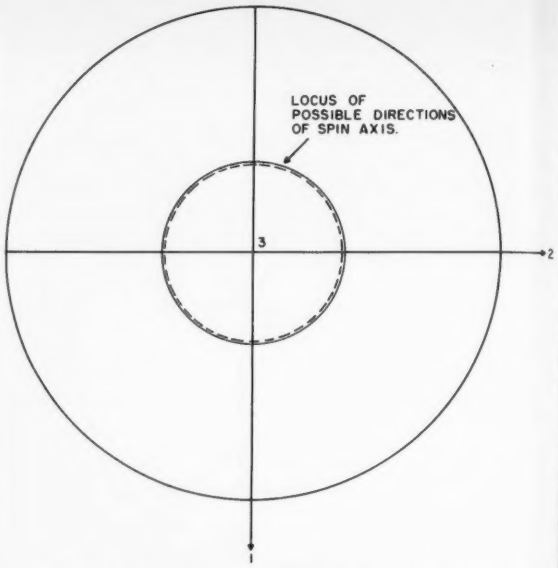


FIGURE 9. *The possible directions of the spin axis, for a given value of* $|\cos\Psi|$.

5.3. The Angle ϕ

ϕ is measured from axis 1 to the upgoing end of the antenna axis, at an instant when the latter passes through the 1-2 plane (sec. 4.2); its possible values range from −180° to +180°. The value of ϕ may be obtained from the times of change of the sign of the correlation. Recall, from section 4.5, that the correlation changes sign at two instants in each half-revolution of the satellite: the first, which can be identified readily since it is also the instant when $|H_a|$ is a maximum, occurs when the antenna axis passes through the 1-2 plane ($\Omega t=0$): the second occurs when it passes through the 1-3 plane. If the observed time delay between the two instants is expressed, in terms of the period of rotation of the satellite, as a phase angle Ωt, then ϕ can be calculated by using (44). After rearrangement, this equation gives

$$\tan\phi=-\cos\Psi\tan\Omega t. \quad (49)$$

Altogether, there are now four possibilities, since for each value of ψ there are two values of ϕ that satisfy (49), and ψ itself has two possible values. The problem now is how to resolve this ambiguity and select the correct combination of ψ and ϕ.

5.4. Resolution of the Ambiguity

If their true values are denoted by the symbols Ψ_t and ϕ_t, the four combinations of Ψ and ϕ that would be possible in the light of the present evidence are the following:

(a)	ψ_t	ϕ_t
(b)	ψ_t	$\phi_t-180°$
(c)	$180°-\psi_t$	$180°-\phi_t$
(d)	$180°-\psi_t$	$-\phi_t$.

Figure 10 illustrates these four possibilities in the analysis of the record of figure 8b; figure 10a shows the possible directions of the spin axis, and figure 10b the corresponding loci of the antenna axis.

What allows this ambiguity to be resolved is the fact that the angle θ is not completely unknown at the outset of the experiment. Thus, in the example, θ is certainly greater than 90°, because the transmitter is located in the northern hemisphere where the lines of magnetic force point downwards, while the waves are traveling upwards. Granted this knowledge, the alternatives (b) and (d) can be eliminated by considering the direction in which the correlation changes sign at $\Omega t=0$; in the example, the sign changes from negative to positive at this instant. Now, since θ is greater than 90°, the sign is governed by the rule on the right-hand side of figure 7b. Applying this rule, it appears that at the instant $\Omega t=0$ the antenna axis is entering the upper right-hand quadrant of the sphere, and hence its locus must be either (a) or (c).

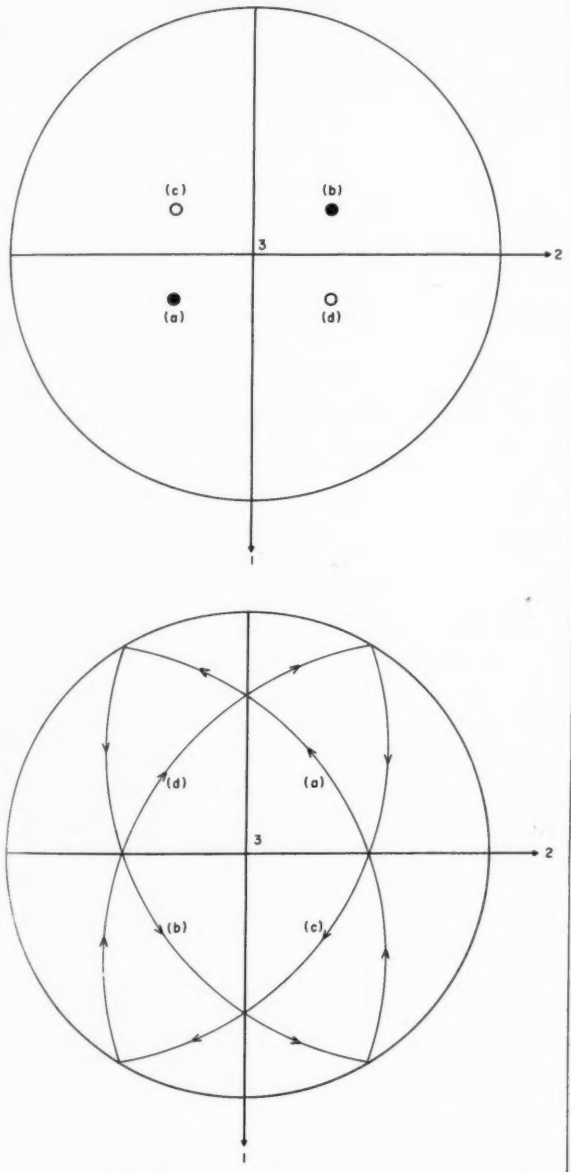


FIGURE 10. *The fourfold ambiguity in the determination of the direction of the spin axis.*

(a) The possible directions of the spin axis, (b) the corresponding loci of the direction of the antenna axis.

The angle θ can be estimated more closely by assuming that the wave normal is directed to the vertical by refraction (see sec. 2.1), and by making use of the known value of the magnetic dip; at a guess, such an estimate would lie within 10° of the true value. If so, then the choice between the two remaining alternatives can be made by considering the variations of the amplitude of the electric signal. The expected variations of $|E_a|$ are calculated for the two possible loci, using the estimated value of θ, and they are compared with the observed variation. The two calculated variations are mutually opposite, so that usually one should fit the record much better than the other, in spite of the inaccuracy of the estimate of θ. Figure 11 shows the comparison for the given example, taking 150° as the estimate of Y. This estimate is in error by 10°, but nevertheless it is clear that (a) is the correct locus. Thus, the ambiguity is resolved.

In practice, the direction of the spin axis in space should be fairly well known from the circumstances of the launching, so that the ambiguity could be resolved immediately, and the chain of reasoning described above would be superfluous.

5.5. The Angle θ

θ is the inclination of the wave normal to the direction of the earth's field, and its possible values lies between 0° and 180°. Now that Ψ and ϕ are known, θ may be found by adjusting its value until the observed and calculated variations of the amplitude of the electric signal agree most closely.

When θ is known, the local electron density N can be calculated from the equations given in section 5.1.

Since θ is estimated from the details of the modulation of the received signals by the spinning motion of the satellite, obviously the depth of this modulation is one of the main factors that control the accuracy of the estimate. Now the depth of the modulation is governed chiefly by the inclination of the spin axis to the wave normal, that is to say, by the value of Ψ. The most favorable condition is that $\Psi\cong 90°$, which gives 100 percent modulation of the magnetic signal; in section 3.4, the absolute direction of the spin axis in space was chosen so as to fulfill this condition when the satellite passes over the transmitter. On the other hand, the least favorable condition is that $\Psi\cong 0°$, where neither the magnetic nor the electric signal is modulated. If this latter condition were to arise, then θ could not

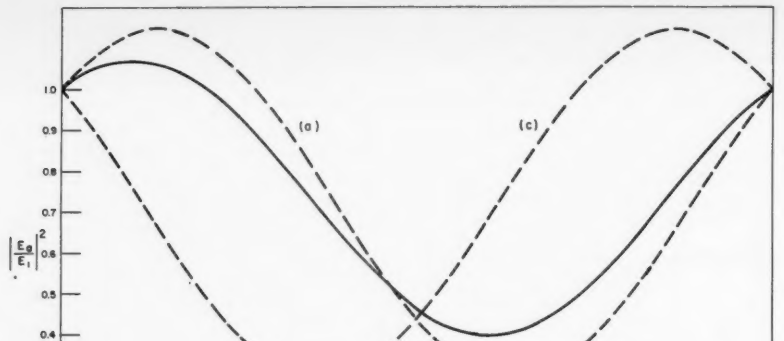


FIGURE 11. *Comparison of the "observed" and calculated variations of the strength of the electric signal; in the calculations, it was assumed that* $\theta = 150°$.

be determined from the records. A theoretical estimate of θ, obtained as described in section 5.4 above, would then have to be used in calculating the electron density N. Note therefore, from (37), that the inferred value of N is roughly proportional to $|\cos\theta|$, so that it should be insensitive to errors in θ around the assumed value of 160°.

6. Integrated VLF Experiment

In the experiment as proposed so far, the satellite measures the least number of quantities that suffice to determine the electron density. Much more could be learned about the ionosphere if, at the same time, measurements were made of certain other quantities; namely, (a) The aspect of the satellite, (b) the local magnetic field, (c) a component of the vlf wave field ($\underline{H}$ or $\underline{E}$) parallel to the spin axis. These measurements will be discussed in turn.

(*a*) *Aspect.* The instantaneous aspect of the satellite could be measured by photoelectric devices which sense the direction of the sun, the moon, or the horizon of the earth. Such measurements would determine the absolute direction of the spin axis in space.

This information, when combined with the record of the amplitude of the magnetic signal H_a, would determine the absolute direction of the wave-normal.

(*b*) *Magnetic field.* Three components of the local magnetic field could be measured by a pair of magnetometers. One, a fluxgate, would measure the component of field along the spin axis. The other, an induction magnetometer, would be set up in the plane transverse to the spin axis; it would make use of the spin of the satellite to induce an alternating voltage that would represent the two components of the magnetic field in the transverse plane. A magnetometer of this type has been used successfully in the lunar probe "Pioneer I" (Sonnett [9]).

From the measurements of the three components, the total local field strength (H_0) can be calculated. These same measurements, when combined with the measurement of aspect, also determine the absolute direction of the field.

Having determined the absolute directions both of the field and of the wave-normal, the angle (θ) between them can be obtained directly. This value is likely to be more accurate than that calculated by the method of section 5, where much use was made of the sign of the correlation between H_a and E_a; this measurement is apt to be sensitive to errors caused by small phase-shifts in the receivers, and by noise also.

It will be recalled that the values of the H_0 and θ are needed for calculating the local electron density from the measured wave admittance (sec. 5.1).

(c) *Component of vlf field along the spin axis.* Finally, following a proposal by Helliwell [10], the satellite could be equipped with a third vlf receiver, connected to an antenna that was alined with the spin axis. The signal received on such an antenna would not be affected by the spinning motion. Its phase (relative to a stable local standard carried in the satellite), and also its amplitude, would be telemetered back to the ground.

The rate of change of phase, when corrected for the change in path length for the telemetry transmission, yields the Doppler frequency shift of the vlf signals as observed at the satellite. This frequency shift is given by

$$\Delta f = (V \operatorname{Cos} \delta)/\lambda \qquad (50)$$

where V is the speed of the satellite along its orbit, δ is the angle between the orbit and the wave-normal, and λ is the wavelength in the medium. Now the orbit is presumed known, and the local direction of the wave normal is measured, so the values of V

and δ are both data. Hence, if Δf is appreciable, λ can be calculated from its observed value. Now

$$\lambda = \lambda_0/\mu \tag{51}$$

where λ_0 is the wavelength in free space. Given λ and λ_0, μ can be calculated. Thus the Doppler shift would provide another measure of the local refractive index, and hence of the electron density; this measurement would check the value calculated from the wave admittance (sec. 5.1).

If the vlf transmissions were modulated in some way, the modulation would be received at the satellite with an appreciable delay, because the group velocity of the whistler mode is rather low. The delay could be measured most easily on this third vlf channel, where the received signals are not modulated in addition by the spin of the satellite. From the observed delay, it is possible to determine the integral of the plasma frequency with respect to height up to the level of the satellite (Helliwell [10]).

The proposal is made that there should be an integrated vlf satellite experiment, incorporating all the features described above. The following information about the ionosphere would be obtained at each point on the orbit: (a) Local strength of earth's magnetic field, (b) local direction of field, (c) local strength of vlf waves, (d) local direction of propagation, (e) local electron density, (f) height integral of plasma frequency.

The changes in these quantities, as the satellite moves along its orbit, would be expected to show both systematic and random components. The systematic components would be related to the bulk properties of the ionosphere, while the random variations would give information on ionospheric irregularities.

7. Discussion

The main proposal made in this paper is for a satellite that would determine local electron density in the ionosphere, by exploring the structure of a vlf wave field set up by a transmitter on the ground. The following sources of difficulty are anticipated:

(a) *Noise.* One limit to the accuracy of the experiment will be set by noise. Judging by previous experience at vlf, there should be no difficulty in making the receivers sufficiently quiet that the limit is set by the external atmospheric noise picked up on the antennas, which is unavoidable. However, the actual level of atmospheric noise within the ionosphere is hard to predict. At the ground, vlf noise is observed to come both from the lower atmosphere, where it is produced by lightning strokes, and also from the ionosphere, where its mode of origin is uncertain (Watts [11]). Within the ionosphere, presumably, the former source of noise is less effective and the latter more so. It is hoped that the vlf satellite experiment already planned by Helliwell [10] will provide data on the noise level in the ionosphere.

(b) *Multiple waves.* The theory has assumed that the vlf field in the ionosphere is that of a single plane wave (sec. 2.1). Thus the experiment would not be expected to work near the conjugate point, where the downcoming direct waves would be mixed with upgoing waves reflected from the base of the ionosphere. Difficulty might be experienced even near the transmitter, if the downcoming wave returned from the opposite hemisphere was at all comparable in strength to the upgoing primary wave. This difficulty could be overcome either by working at a higher frequency, which would be more heavily absorbed, or by modulating the transmissions in such a way that the direct signal and the echo never overlap one another; for 15.5 kc signals from the station NSS, the time of travel over the whistler route and back would be about 1 sec (Helliwell and Gehrels [12]), so that a suitable form of modulation would be a ½ sec pulse repeated every 2 sec. A potentially more serious problem is presented by ionospheric irregularities, which would distort the upgoing wavefronts. If the curvature of the wavefronts was appreciable in terms of a wavelength, then the plane-wave theory would not apply. The significance of the measurement of wave admittance in these circumstances has not yet been investigated.

(c) *Effect of the medium on the electric dipole antenna.* There is doubt as to how well the electric dipole antenna will perform when immersed in the anisotropic plasma of the ionosphere. The design of antenna that was described in section 3.2. is intended to minimize the adverse effects of the medium. If such effects persist nevertheless, their presence could be detected from the fact that they will almost certainly cause the variation of $|E_a|^2$ with time to depart from the pure sinusoid predicted by the theory (sec. 4.4).

(d) *Disturbance of the medium by the satellite.* There is a chance that the satellite, by virtue of its rapid orbital motion, may disturb the medium seriously and thus alter the distribution of magnetic and electric fields in its immediate neighbourhood. The purely hydrodynamic disturbance is likely to be small, and confined to the region behind the satellite (sec. 3.2). However, Kraus and Watson [13] have pointed out that the satellite, in its passage through the ionosphere, acquires an electrostatic charge, and that the field of this charge should produce a hydromagnetic disturbance ahead of the satellite. It is uncertain how serious this effect will be. In this connection, it is worth noting that instruments have been developed to measure the state of charge of a satellite (Krassovsky [14]), and doubtless one could be arranged to control some discharging device so as to maintain a state of electrical neutrality, and thus eliminate the effect.

(e) *Incorrect spinning motion.* Finally, there is the possibility of failure to establish the desired spinning motion: the satellite might be launched either with no spin, or spinning too slowly, or again the spin might be fast enough but accompanied by a spurious wobble that would take some time to

suppress (see secs. 3.4 and 3.5). If the spin was non-existent or very slow, then the experiment would be spoiled completely. A wobble would not affect the measurement of wave admittance, since the spin would still bring the antenna axis through the plane of the wavefront (see sec. 5.1), but it would spoil the measurement of θ. In such a case the electron density could still be calculated by using a theoretical estimate of θ (sec. 5.4), and this procedure would probably not entail much error (sec. 5.5), unless θ was varying widely and at random along the orbit due to the effects of ionospheric irregularities; this last condition would be obvious from the records of signal amplitude. The presence of a component of wobble in the spin motion could be detected most easily if the satellite carried some device for measuring aspect.

In view of these several difficulties that may beset the determination of local electron density by the wave admittance method, it is clear that this experiment should not be performed in isolation, but only as part of an integrated vlf satellite experiment such as that proposed in section 6, where some of the measurements would provide cross-checks on others.

The author is indebted for advice from C. G. Little, R. M. Gallet, and R. S. Lawrence of this Laboratory, from J. W. Warwick of the High Altitude Observatory of the University of Colorado, and from R. A. Helliwell of Stanford University. The proposal for an integrated vlf satellite experiment arose out of a discussion with Dr. Helliwell.

8. References

[1] W. C. Hoffman, Report on the symposium on the use of space vehicles at the fall 1958 URSI meeting (to be published in J. Geophys. Research).
[2] L. R. O. Storey, Phil. Trans. Roy. Soc London **246**, 113 (1953).
[3] H. G. Booker, Cornell Univ. School of Elec. Eng., Tech. Rept. No. 1 (March 1950).
[4] R. A. Helliwell, Stanford Univ., Radio Propagation Lab., Rept. No. AFCRC–TR–56–189.
[5] J. C. Seddon, Reported in reference [1].
[6] R. N. Bracewell and O. K. Garriott, Nature **182**, 760 (1958).
[7] A. E. H. Tutton, Crystallography and practical crystal measurement vol. 1, p. 62 (Macmillan & Co., Ltd., London, England, 1922).
[8] C. W. Bunn, Chemical crystallography, p. 29 (Oxford Univ. Press, Oxford, England, 1945).
[9] C. P. Sonett, Proc. of Lunar and Planetary Exploration Colloquium, N. Am. Aviation (to be published).
[10] R. A. Helliwell, Reported in reference [1].
[11] J. M. Watts, J. Geophys. Research **62**, 199 (1957).
[12] R. A. Helliwell and E. Gehrels, Proc. IRE **46**, 785 (1958).
[13] L. Kraus and K. M. Watson, Phys. Fluids **1**, 480 (1958).
[14] V. I. Krassovsky, Proc. IRE **47**, 289 (1959).

BOULDER, COLO. (Paper 63D3–30)

Selected Abstracts

IGY observations of *F*-layer scatter in the Far East, R. Bateman, J. W. Finney, E. K. Smith, L. H. Tveten, and J. M. Watts, *J. Geophys. Research*, **64**, *403 (Apr. 1959)*.

Peculiar signal enhancements observed during transmissions at 36 to 50 Mc between the Philippines and Okinawa appear to represent *F*-layer scatter. These signals are observed nightly for periods of several hours during the months of September and October. Pulse tests indicate *F*-layer heights for these signals. Considerable pulse broadening is observed and the signals generally arrive from somewhat off the great circle path.

Models of the atmospheric radio refractive index, B. R. Bean and G. D. Thayer, *Proc. IRE*, **47**, *740 (May 1959)*.

This paper introduces two models of atmospheric radio refractive index which can be used to predict refraction effects from the value of the refractive index at the transmitting point. Both models offer considerable improvement over the usual four-thirds earth model, particularly for applications at long distances and high elevations in the atmosphere. Further, both models may be adjusted to represent mean conditions at different times of year and in different geographical locations. A new method of predicting radio-ray refraction at very low initial elevation angles is introduced which utilizes both the initial value and the initial height-gradient of the refractive index over roughly the first 100 meters above the earth's surface.

On the theory of fading properties of a fluctuating signal imposed on a constant signal, H. Bremmer, *NBS Circ. 599 (May 1959)*.

This paper deals with a theoretical investigation of the fading properties of a signal composed of a fluctuating contribution and another steady contribution with fixed amplitude and phase. It is assumed that the central-limit theorem may be applied to two proper quantities describing the fading signal as a quasi-monochromatic function of the time. The results are applicable to any autocorrelation function for the fluctuating contribution. The first part of the paper (secs. 1 to 16) is mainly restricted to the idealized case in which any two components of the fluctuating part of the complete signal that are in quadrature with respect to their phase do have identical statistical properties; the fluctuating part is then termed a "random" signal. This idealized case is shown to constitute but an approximation if applied to the fluctuating field due to first order scattering in a turbulent atmosphere. Therefore, in the second part of the paper (secs. 17 to 26) the theory has been extended to fluctuating contributions (then termed "quasi-random" contributions) not satisfying the above condition of isotropy. All results then depend on two complex correlation functions $a(\tau)$ and $b(\tau)$ instead of on the single function $a(\tau)$ governing the simplified theory. In contrast to $a(\tau)$, the function $b(\tau)$ does not exclusively depend on the energy spectrum of the fluctuating contribution. The fading properties investigated for the composed signal are the distribution functions of both the amplitude and phase, as well as the average number of crossings of each of them (per unit time interval) through any given level. The complicated general formulas reduce to simple expressions in the two extreme cases of (a) absence of the steady signal (e.g., tropospheric scatter propagation to distances far beyond the transmitter's horizon), and (b) predominance of the steady signal (e.g., line-of-sight propagation to distances well within the horizon). The first limiting case leads, when neglecting the influence of $b(\tau)$ (then being very small), to the well-known Rayleigh distribution for the amplitude, to a homogeneous distribution for the phase, and to a fading rapidity (defined as the average number of crossings per unit time interval through the level most frequently passed) which is for the amplitude faster than for the phase by a factor of 3.04. The second limiting case amounts to normal distributions for both amplitude and phase, and to fading rapidities of these quantities which are only identical insofar as the different behavior of the in-phase and in-quadrature component (with respect to the steady signal) does have no numerical importance. On the other hand, measurements of the difference in the amplitude-fading rapidity and the phase-fading rapidity will reveal the effect of the asymmetry with respect to the two mentioned components of the scattered signal.

Sporadic *E* at VHF in the USA, R. M. Davis, Jr., E. K. Smith, and C. D. Ellyett, *Proc. IRE* **47**, *762 (May 1959)*.

An analysis is made of sporadic-*E* propagation observed on the Cedar Rapids to Sterling path during the four years, 1952 to 1955. VHF transmissions over this path at frequencies of 27.775 Mc and 49.8 Mc provided the *Es* data. At this latitude, the summer months receive the preponderance of sporadic *E*. Diurnally, sporadic-*E* occurrence tends to favor the daytime and evening hour with peaks of incidence about 100 and 1800. No variation with sunspot number has been discerned. As a rule, higher *Es* signal intensities are recorded at 28 Mc than at 50 Mc. Cumulative distributions of signal intensity are presented for the two frequencies. A relationship is found between the frequency dependence of *Es* signal intensities ≥ -70 db, relative to inverse distance, and the distribution of *fEs* values at Washington, D.C. The relationship promises to be useful in the prediction of *Es* signal intensities on a worldwide basis. An inverse correspondence is shown between sporadic-*E* occurrence and geomagnetic activity. The correspondence holds only over selected time intervals. The frequency dependence of received power under sporadic-*E* conditions is different from that during normal scatter. The median frequency exponent is two or more times as large for sporadic *E*, and the exponents cover a much wider range of values. This is tentatively explained by considering the *Es* region to be composed of patches of intense nonuniform ionization, an hypothesis previously used to explain vertical-incidence data.

Diurnal change of ionospheric heights deduced from phase velocity measurements at VLF, James R. Wait, *Proc. IRE* **47**, *998 (May 1959)*.

The phase shift of the carrier of GBR (Rugby) recorded in Massachusetts and New Zealand is interpreted in terms of a diurnal change of ionospheric reflecting height of about 17 km. The results provide further confirmation of the waveguide model for terrestrial propagation of radio waves at very low frequencies.

On the theory of reflection from a wire grid parallel to an interface between homogeneous media (II), James R. Wait, Appl. Sci. Research, sec. B **7**, 355 (1959).

A previous analysis is extended to include coupling between the TM and TE modes. This can be important when the incidence and polarization are both arbitrary.

Transmission loss in radio propagation II, K. A. Norton, *TN12 (PB151371) $3.00.*

In an earlier report with this title the concept of transmission loss was defined and its advantages explained. In this report a survey is made of the transmission losses expected for a wide range of conditions, i.e., for distances from 10 to 10,000 statute miles; for radiofrequencies from 10 kc to 100,000 Mc; for vertical or horizontal polarization; for ground waves, ionospheric waves, and tropospheric waves; over sea water or over land which may be either rough or smooth; and for various geographical and climatological regions.

Parts of this Technical Note have been published in the following references:

(1) K. A. Norton, Low and medium frequency radio propagation, Proc. Intern. Congr. Propagation of Radio Waves at Liége, Belgium, Oct. 1958 (to be published by the Academic Press).
(2) K. A. Norton, System loss in radio wave propagation, J. Research NBS **63D,** 53 (July 1959).
(3) K. A. Norton, System loss in radio wave propagation, Letter to the Editor, Proc. IRE **47,** 1661 (Sept. 1959).

This Technical Note contains, in addition, a discussion of transmission losses expected in propagation via the regular ionospheric layers and via ionospheric scatter at frequencies in the range from 1 Mc to 100 Mc. Also, one of the appendices is devoted to a physical theory of ionospheric scatter propagation.

Technical considerations leading to an optimum allocation of radiofrequencies in the band 25 to 60 Mc, K. A. Norton, *TN13 (PB151372) $2.50.*

The use of the concept of transmission loss is first explained as it relates to the problem of determining the mutual interference to be expected between services occupying the same or adjacent portions of the frequency spectrum. It is then shown that services operating on the same radiofrequencies in this particular band are subject at times to severe mutual interference due to (a) transmission by way of the regular *F*2 layer of the ionosphere, or (b) transmission by way of sporadic *E* ionization of the ionosphere.

It is shown that *F*2 layer interference is the most severe in terms of its intensity. However, this kind of interference decreases systematically in its duration of occurrence as the radiofrequency is increased, and also varies to a substantial degree geographically, diurnally, seasonally, and with sunspot cycle.

Sporadic *E* interference may also be quite intense, but the probability of its occurrence (or the percentage of time that it exists) decreases rapidly with increasing intensity of the interfering signals and with increasing frequency. The geographical, diurnal, seasonal, and sunspot cycle variations of this mode of propagation are not as well understood or as predictable as *F*2 interference.

Other modes of propagation of potential interference are ionospheric and tropospheric forward scatter, but the interference arising from these modes is usually much less intense and can often be avoided by careful planning of the use of this portion of the spectrum in terms of station separations and the use of directive antennas.

It is shown in an appendix that the power required to provide a given service may either increase or decrease with increasing frequency, depending largely on the kind of antennas used. Thus, mobile services which normally use nondirectional antennas at both ends of the circuit require more power for a given range the higher the frequency. On the other hand, point-to-point services may use large directional arrays at both ends of the circuit and, for a given size, these arrays have an increased gain the higher the radiofrequency; thus the point-to-point services may often operate more efficiently at the highest frequencies free from the effects of atmospheric absorption, i.e., above 2000 Mc and below, say, 20,000 Mc.

Based on the above conclusions, the following allocation procedures are recommended by the author:

a. Since the new High Power Fixed Service (RPIS) provides a kind of communication facility which cannot readily be duplicated elsewhere in the entire radio spectrum, it should be favored relative to other services occupying this particular portion of the spectrum.

b. Because of the ease of jamming by means of *F*2 propagation, all services likely to be jammed should use frequencies sufficiently high so that *F*2 propagation will not occur during most, if not all, of the time.

c. High Power Fixed Services likely to be jammed should be allocated frequencies sufficiently high so that *F*2 propagation is not feasible even at sunspot maximum and geographically separated (by, say, 1,500 miles or more) from other users of these same frequencies so that other modes of propagation do not cause interference.

d. High Power Fixed Services unlikely to be jammed may be allocated cleared channels anywhere within the 25- to 60-Mc range. Such cleared channels would necessarily have to be negotiated and cleared internationally.

e. Since the Land Mobile Services are evidently able to operate as at present under conditions of fairly severe sporadic interference such as that caused by tropospheric and sporadic *E* propagation, it seems likely that many more such stations can be allocated to individual channels in this band before the loss in service suffered by those stations already on the channels becomes greater than the additional service provided by the newly allocated stations. Only when such a balance is struck can it truly be said that efficient use is being made of this portion of the spectrum.

f. By the use of higher transmitter-carrier-frequency and receiver-oscillator-frequency stabilities, narrower range modulation schemes, and improved receivers, more channels may be made available per megacycle of this valuable spectrum space.'

g. Sufficiently high power should be used by all services so that noise does not limit the communications at any time in the case of point-to-point services or at any location within normal operational range in the case of mobile or broadcast services.

h. Many land mobile services can operate satisfactorily in higher frequency portions of the spectrum and, especially if reliable, continuous service is essential, should seek channels in these higher frequency bands.

Guiding of electromagnetic waves by uniformly rough surfaces (Parts I and II), James R. Wait. *Paper published in Proceedings of URSI International Symposium on Electromagnetic Theory, University of Toronto, June 15–20, 1959. (Proceedings are available from the Institute of Radio Engineers, 1 East 79th St., New York 21, N.Y. Rates: $16.00 a copy for non-I.R.E. members, $12.00 for I.R.E. members and libraries.)*

Part I. A simple derivation is given for the reflection of electromagnetic waves from a perfectly conducting plane surface which has a uniform distribution of hemispherical bosses whose electrical constants are arbitrary. The spacing between the centers of the bosses is taken to be small which is the justification for neglecting the incoherent radiation. An approximate boundary condition is developed which must be satisfied in an average sense by the tangential fields on the reference plane.

The excitation of surface waves on the rough surface is then discussed. It is indicated that to a first order, a rough surface of the kind described here possesses an inductive surface reactance and will support a trapped wave. The effect of finite conductivity of the bosses is to exponentially damp this trapped wave.

Part II. Using the model of a single rough surface described in Part I, the influence of curvature is considered. The starting point is a residue-series solution for a vertical dipole over a sphere with an arbitrary surface impedance. It is shown that the curvature has a profound influence on the nature of the surface wave, although it uniformly approaches the conditions for a plane boundary as the radius of curvature approaches infinity.

The propagation between the space bounded by two parallel uniformly rough surfaces is also considered. This is treated first for plane boundaries and then for concentric spherical boundaries. The latter model is useful in explaining certain experimental data on the terrestrial propagation of VLF radio waves.

Other surface-wave papers in the symposium are authored by H. Bremmer, J. Brown, A. L. Cullen, L. B. Felsen, B. Friedman, K. Furutsu, G. Goubau, A. Hessel, A. E. Karbowiak, H. M. Barlow, A. A. Oliner, G. Piefke, S. A. Schelkunoff, and F. J. Zucker.

Other NBS Publications*

Journal of Research, Section 63B, No. 2, October–December 1959. 75 cents.

Applications of a theorem on partitioned matrices. E. V. Haynsworth.

Capacity requirement of a mail sorting device: II. A. J. Goldman.

Analytic comparison of suggested configurations for automatic mail sorting equipment. B. K. Bender and A. J. Goldman.

New method of solution for unretarded satellite orbits. John P. Vinti.

Effect of sudden water release on the reservoir free outflow hydrograph. Vujica M. Yevdjevich.

Uniform asymptotic expansions for Weber parabolic cylinder functions of large orders. F. W. J. Olver.

Journal of Research, Section 63C, No. 2, October–December 1959. 75 cents.

Conical coaxial capacitors and their advantages. M. C. Selby.

A photoelectric followup and recording system, and its application to remote observations of the beam in high precision balances. H. A. Bowman and L. B. Macurdy.

A stroboscopic vibration analyzer. S. Edelman, R. Brooks, S. Saito, E. Jones, and E. R. Smith.

Evaluation of lens distortion by the inverse nodal slide. Francis E. Washer and Walter R. Darling.

Evaluation of lens distortion by the modified goniometric method. Francis E. Washer and Walter R. Darling.

Proposed criteria for defining load failure of beams, floors, and roof constructions during fire tests. J. V. Ryan and A. F. Robertson.

Conductive flooring for hospital operating rooms. Thomas H. Boone, Francis L. Hermach, Edgar H. MacArthur, and Rita C. McAuliff.

Measurement of the aging of rubber vulcanizates. J. Mandel, F. L. Roth, M. N. Steel, and R. D. Stiehler.

Thermodynamic properties of helium at low temperatures and high blood pressures, D. B. Mann and R. B. Stewart, TN8 (PB151367) $1.25.

Transmission loss in radio propagation—II, K. A. Norton, TN12 (PB151371) $3.00. (See abstracts above.)

Technical considerations leading to an optimum allocation of radiofrequencies in the band 25 to 60 Mc, K. A. Norton, $2.50. (See abstracts above.)

The role of government research laboratories, A. V. Astin, Elec. Engr. **78**, No. 7, 738 (1959).

The ampere, F. B. Silsbee, Proc. IRE **47**, No. 5, 643 (1959).

The basis of our measuring system, A. G. McNish, Proc. IRE **47**, No. 5, 636 (1959).

Reflectors for a microwave Fabry-Perot interferometer, W. Culshaw, IRE Trans. on Microwave Theory and Techniques MTT–**7**, No. 2, 221 (1959).

The research frontier, I. C. Schoonover, Saturday Rev., p. 58 (May 2, 1959).

The nature, cause, and effect of the porosity in electrodeposits, III. Microscopic detection of porosity, F. Ogburn and D. W. Ernst, Plating **46**, 831 (1959).

Paper test data from pendulum and inertialess testers, R. J. Capott, S. B. Newman, and J. Mandel, TAPPI **42**, No. 6, 480 (1959).

Convexity of the filed of a linear transformation, A. J. Goldman and M. Marcus, Can. Math. Bull. **2**, No. 1, 15 (1959).

Factors affecting the accuracy of distortion measurements made on the nodal slide optical bench, F. E. Washer and W. R. Darling, J. Opt. Soc. Am. **49**, No. 6, 517 (1959).

Mechanism of contraction in the muscle fiber-ATP system, L. Mandelkern, A. S. Posner, A. F. Diorio, and K. Laki, Proc. Natl. Acad. Sci. **45**, No. 6, 814 (1959).

Microwave spectrum of methyl germane, V. W. Laurie, J. Chem. Phys. **30**, No. 5, 1210 (1959).

IGY instruction manual. Part 1: World days and communications, A. H. Shapley, Annals of IGY **7**, Pt. 1, 1 (Pergamon Press, New York, N.Y., 1959).

Measurement of ozone in terms of its optical absorption, R. Stair, Advances in Chem. Series of the Am. Chem. Soc., No. 21, 269 (1959).

On the perturbation of the vibrational equilibrium distribution of reactant molecules by chemical reactions, K. E. Shuler, 7th Symp. (Intern.) on Combustion, London and Oxford, Aug. 28 to Sept. 3, 1958. Combustion Inst., p. 87 (Butterworths Sci. Pub., London, England, 1958).

The calculation of the field in a homogeneous conductor with a wavy interface, J. R. Wait, Proc. IRE **47**, 1155 (1959).

Rack for standard resistors, P. H. Lowrie, Jr., Rev. Sci. Instr. **30**, No. 4, 291 (1959).

Lower bounds for eigenvalues with application to the helium atom, N. W. Bazley, Proc. Natl. Acad. Sci. **45**, No. 6, 850 (1959).

Graphical diagnosis of interlaboratory test results, W. J. Youden, Ind. Qual. Control **XV**, No. 11, 1 (1959).

Spectroscopic evidence for triatomic nitrogen in solids at very low termperature, M. Peyron, E. M. Horl, H. W. Brown, and H. P. Broida, J. Chem. Phys. **30**, No. 5, 1304 (1959).

The evaluation of small color differences: Part I. Visual observations, J. C. Richmond and W. N. Harrison, Am. Ceram. Bul. **38**, No. 6, 292 (1959).

Radiation from a small loop immersed in a semi-infinite conducting medium, James R. Wait, *Can. J.* Phys. **37**, 672 (1959).

**Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (Order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

INDEX TO VOLUME 63D

July to December 1959

A

Page

Agy, V., K. Davies, Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Airglow and aurora, origin of [OI] 5577 ... 15
Antenna gain, path, in an exponential atmosphere ... 273
Antennas, slotted-cylinder, pattern synthesis ... 303
Atmosphere, exponential reference, Central Radio Propagation Laboratory ... 287
Atmosphere, path antenna gain in an exponential ... 315
Atmosphere, upper, excitation mechanisms of the oxygen 5577 emission ... 273
Atmospheric horizontal inhomogeneity, effect of upon ray tracing ... 319
Aurora, a monochromatic low-latitude ... 297
Auroral and subauroral zones, [OI] 5577 in, comparison of absolute intensities ... 19

B

Bean, B. R., Climatology of ground-based radio ducts ... 29
—, *B. A. Cahoon,* Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
—, *J. D. Horn,* Radio-refractive-index climate near the ground ... 259
—, *L. P. Riggs, J. D. Horn,* Synoptic study of the vertical distribution of the radio refractive index ... 249
—, —, Synoptic variation of the radio refractive index ... 91
—, *G. D. Thayer,* Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Bremmer, H., Mode expansion in the low-frequency range for propagation through a curved stratified atmosphere ... 75
Broadcasting station, worldwide standard-frequency, power requirements and choice of an optimum frequency ... 35
Budden, K. G., Effect of small irregularities on the constitutive relations for the ionosphere ... 135

C

Cahoon, B. A., B. R. Bean, Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Climate, radio-refractive-index, near the ground ... 259
Climatology of ground-based radio ducts ... 29
Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
Conda, A. M., J. R. Wait, Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181
Correlation of solar noise fluctuations in harmonically related bands ... 293

D

Davies, K., V. Agy, Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Decker, M. T., Transmission and reflection by a parallel wire grid ... 87
Density, local electron, method for measuring from an artificial satellite ... 312
Devaney, T. E., A. N. Smith, Fields in electrically short ground systems ... 175
Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181

E

Effect of atmospheric horizontal inhomogeneity upon ray tracing ... 287
Effect of small irregularities on the constitutive relations for the ionosphere ... 135
Electrically short ground systems, fields ... 175
Electromagnetic waves, diffraction of, by smooth obstacles for grazing angles ... 181
Electron density, local, method for measuring from an artificial satellite ... 325
Ellyett, C., J. M. Watts, Stratification in the lower ionosphere ... 117
Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319
Exponential atmosphere, path antenna gain ... 273
Exponential reference atmosphere, Central Radio Propagation Laboratory ... 315

F

Fields in electrically short ground systems ... 175
Fluctuations in harmonically related bands, correlation of solar noise ... 293
Frequency and power requirements for a worldwide standard-frequency broadcasting station ... 35

G

Gallet, R. M., R. A. Helliwell, Origin of "very-low-frequency emissions" ... 21
Gatterer, D. M., Preliminary results of the National Bureau of Standards radio and ionospheric observations during the International Geophysical Year ... 1
Grazing angles, diffraction of electromagnetic waves by smooth obstacles ... 181
Grazing angles, study at 1,046 megacycles per second of the reflection coefficient of irregular terrain ... 235
Grid, parallel wire, transmission and reflection ... 87
Ground-based radio ducts, climatology ... 29
Ground systems, electrically short, fields ... 175

H

Hartman, W. J., R. E. Wilkerson, Path antenna gain in an exponential atmosphere ... 273
Helliwell, R. A., R. M. Gallet, Origin of "very-low-frequency emissions" ... 21
Horn, J. T., B. R. Bean, Radio refractive-index climate near the ground ... 259
—, —, *L. P. Riggs,* Synoptic study of the vertical distribution of the radio refractive index ... 349
Householder, J., J. R. Wait, Pattern synthesis for slotted-cylinder antennas ... 303

I

Index, radio refractive, synoptic study of the vertical distribution ... 249
Intensities, absolute, of [OI] 5577 in the auroral and subauroral zones, comparison ... 19
International Geophysical Year, radio and ionospheric observations ... 1
Ionosphere, effect of small irregularities on the constitutive relations ... 135
Ionosphere, lower, stratification in ... 117
Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence ... 151
Irregularities, small, effect of on the constitutive relations for the ionosphere ... 135
Irregularities, tropospheric, radio-wave scattering ... 205

J

Janes, H. B., M. C. Thompson, Jr., Measurements of phase stability over a low-level tropospheric path ... 45
Jean, A. G., W. L. Taylor, Very-low-frequency radiation spectra of lightning discharges ... 199

L

Lightning discharges, vlf radiation spectra of ... 199
Low-frequency propagation paths in arctic areas ... 99
Low-latitude aurora, a monochromatic ... 297

M

Maloney, L. J., R. E. McGavin, Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Marovich, E., F. E. Roach, A monochromatic low-latitude aurora ... 297
—, —, *J. W. McCaulley,* Origin of [OI] 5577 in the airglow and the aurora ... 15
Maxwell, E. L., A. D. Watt, Low-frequency propagation paths in arctic areas ... 99
McCaulley, J. W., F. E. Roach, E. Marovich, Origin of [OI] 5577 in the airglow and the aurora ... 15
—, —, *C. M. Purdy,* Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
McGavin, R. E., L. J. Maloney, Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Measurements of phase stability over a low-level tropospheric path ... 45
Method for measuring local electron density from an artificial satellite ... 325
Mode expansion in the low-frequency range for propagation through a curved stratified atmosphere ... 75
Monochromatic low-latitude aurora ... 297

N

Norton, K. A., System loss in radio wave propagation ... 53

O

Origin of [OI] 5577 in the airglow and the aurora ... 15
Origin of "very-low-frequency emissions" ... 21
Oxygen 5577 emission in the upper atmosphere, excitation mechanisms ... 319

P

Path antenna gain in an exponential atmosphere ... 273
Pattern synthesis for slotted-cylinder antennas ... 303
Phase stability, measurements of, over a low-level tropospheric path ... 45
Plush, R. W., A. D. Watt, Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 35
Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 35
Preliminary results of the National Bureau of Standards radio and ionospheric observations during the International Geophysical Year ... 1
Propagation, low-frequency, in arctic areas ... 99
Propagation through a curved stratified atmosphere, mode expansion in the low-frequency range ... 75
Purdy, C. M., F. E. Roach, J. W. McCaulley, Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19

R

Page

Radiation spectra, vlf, of lightning discharges ... 199
Radio and ionospheric observations during the International Geophysical Year ... 1
Radio ducts, ground-based, climatology ... 29
Radio-refractive-index climate near the ground ... 259
Radio refractive index, synoptic study of the vertical distribution ... 249
Radio refractive index, synoptic variation ... 91
Radio-wave propagation, system loss ... 53
Radio-wave scattering by tropospheric irregularities ... 205
Ray tracing, effect of atmospheric horizontal inhomogeneity ... 287
Reflection and transmission by a parallel wire grid ... 87
Reflection coefficient of irregular terrain at grazing angles, study at 1,046 megacycles per second ... 235
Refractive index, radio, synoptic variation ... 91
Riggs, L. P., B. R. Bean, Synoptic variation of the radio refractive index ... 91
—, —, *J. D. Horn,* Synoptic study of the vertical distribution of the radio refractive index ... 249
Roach, F. E., E. Marovich, A monochromatic low-latitude aurora ... 297
—, *J. W. McCaulley, E. Marovich,* Origin of [OI] 5577 in the airglow and the aurora ... 15
—, —, *C. M. Purdy,* Comparison of absolute intensities of [OI] 5577 in the auroral and subauroral zones ... 19
—, *E. Tandberg-Hanssen,* Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319

S

Satellite, artificial, method for measuring local electron density ... 325
Slotted-cylinder antennas, pattern synthesis ... 303
Smith, A. N., T. E. Devaney, Fields in electrically short ground systems ... 175
Smooth-obstacle diffraction of electromagnetic waves for grazing angles ... 181
Spectra, vlf radiation, of lightning discharges ... 199
Stability, phase, measurements over a low-level tropospheric path ... 45
Standard-frequency broadcasting station, worldwide, the power requirements and choice of an optimum frequency ... 35
Storey, L. R. O., Correlation of solar noise fluctuations in harmonically related bands ... 293
—, Method for measuring local electron density from an artificial satellite ... 325
Stratification in the lower ionosphere ... 117
Study at 1,046 megacycles per second of the reflection coefficient of irregular terrain at grazing angles ... 235
Sweep-frequency pulse technique at oblique incidence, ionospheric investigations using ... 151
Synoptic study of the vertical distribution of the radio refractive index ... 249
Synoptic variation of the radio refractive index ... 91
Synthesis for slotted-cylinder antennas ... 303
System loss in radio wave propagation ... 53

T

Tandberg-Hanssen, E., F. E. Roach, Excitation mechanisms of the oxygen 5577 emission in the upper atmosphere ... 319
Taylor, W. L., A. G. Jean, Very-low-frequency radiation spectra of lightning discharges ... 199
Terrain, irregular, study at 1,046 megacycles per second of the reflection coefficient at grazing angles ... 235
Thayer, G. D., B. R. Bean, Central Radio Propagation Laboratory exponential reference atmosphere ... 315
Thompson, M. C., Jr., H. B. Janes, Measurements of phase stability over a low-level tropospheric path ... 45
Transmission and reflection by a parallel wire grid ... 87
Tropospheric irregularities, radio-wave scattering ... 205
Tropospheric path, low-level, measurements of phase stability ... 45

V

Variation, synoptic, of the radio refractive index ... 91
Very-low-frequency emissions, origin of ... 21
Very-low-frequency radiation spectra of lightning discharges ... 199

W

Wait, J. R., A. M. Conda, Diffraction of electromagnetic waves by smooth obstacles for grazing angles ... 181
—, *J. Householder,* Pattern synthesis for slotted-cylinder antennas ... 303
Watt, A. D., E. L. Maxwell, Low-frequency propagation paths in arctic areas ... 99
—, *R. W. Plush,* Power requirements and choice of an optimum frequency for a worldwide standard-frequency broadcasting station ... 3
Watts, J. M., C. Ellyett, Stratification in the lower ionosphere ... 117
Wheelon, A. D., Radio-wave scattering by tropospheric irregularities ... 205
Wilkerson, R. E., W. J. Hartman, Path antenna gain in an exponential atmosphere ... 273
Worldwide standard-frequency broadcasting station, the power requirements and choice of an optimum frequency ... 35